FACTORIAL ANALYSIS
FOR NON-MATHEMATICIANS

by

C. J. ADCOCK
M.A. (N.Z.), Ph.D. (Lond.)
Senior Lecturer in Psychology
Victoria University College, Wellington, N.Z.

MELBOURNE UNIVERSITY PRESS

Many students of psychology neglect factor analysis because the current texts frighten them by their mathematical complexity. This book sets out to cover the essential processes of factor analysis in such a way as to make clear the underlying logic in simple language. The basic principles are sufficiently simple, and there is no reason why students with the most elementary mathematical background should not sufficiently understand them to make some use of them in research and to be critically appreciative of research done by others.

17/6
Net

Printed in Australia

DAS GUPTA

~~9024~~

~~6656~~
~~9021~~ SIE.

POST-GRADUATE BASIC TRAINING COLLEGE * Banipur, Po. Balgachi, 24-Pgs. *
सत्यमेव जयते

FACTORIAL ANALYSIS
FOR
NON-MATHEMATICIANS

POST-GRADUATE BASIC TRAINING COLLEGE * Banipur, Po. Salgachi, 24-Pgs.

FACTORIAL ANALYSIS

FOR NON-MATHEMATICIANS

by

C. J. ADCOCK

M.A. (N.Z.), Ph.D. (Lond.)

Senior Lecturer in Psychology
Victoria University College, Wellington, N.Z.

MELBOURNE UNIVERSITY PRESS

First published 1954
Printed and bound in Australia by
*Melbourne University Press, Carlton, N.*3*, Victoria*
Registered in Australia for transmission by post as a book

6656

[illegible] West Bengal
[illegible] 7.1.94
[illegible] 7671

London and New York: Cambridge University Press

CONTENTS

INTRODUCTION 7

I. THE LOGIC OF FACTOR ANALYSIS 8

II. CORRELATION 10
- Perfect Positive Correlation 11
- Perfect Negative Correlation 11
- Zero Correlation 12
- Intermediate Positive Correlation 14
- Arithmetical Calculation 15
- Practical Methods 17
- Tetrachoric Correlation 18

III. FACTORS 19
- Correlation Basis 19
- Calculating Factor Loadings 22
- More Than One Factor 23

IV. THE PRINCIPLE OF ROTATION 29
- Simple Structure 30
- Spearman and *g* 35

V. FINDING SIMPLE STRUCTURE 38
- Rotational Methods 38
- Single Plane Method of Rotation 38
- Method of Extended Vectors 42

VI. CLUSTER—DIRECTED ANALYSIS 48
- Selecting the Clusters 49
- Finding the Factors 53

VII. COMBINING CENTROID AND GROUP ANALYSIS 68
- Merits of the Method 76

VIII. INTERPRETING THE RESULTS 77

IX. THE ROLE OF FACTOR ANALYSIS 81
Less Bias of Popular Opinion.. 82
Economy of Tests 83
Validation of Tests 83
The Status of Factors 83

APPENDIX—Basing Factors on Correlation Clusters 85
Formulas for Correction Matrix 87

FURTHER READING 88

INTRODUCTION

Many students of psychology neglect factor analysis because the current texts frighten them by their mathematical complexity. I am not familiar with any text which covers the essentials of factor analysis in simple, non-mathematical terms. There seems to be a need for a book which will enable the student to get a quick insight into what is involved without labouring through any mathematical equations.

The present book sets out to cover the essential processes of factor analysis in such a way as to make clear the underlying logic in simple language. To the mathematician it may seem clumsy and ineffective, but neatness of mathematical statement is no help to those not schooled in its use. The basic processes are sufficiently simple, and there is no reason why students with the most elementary mathematical background should not sufficiently understand them to make some use of them in research and to be critically appreciative of research done by others.

The technique developed within psychology and its chief exponents are still psychologists. It would, therefore, be a great pity if mathematical refinement should hinder its use or discredit the results. It is hoped that this present effort will help to bridge the gap between the experts who are continually adding to the flexibility and effectiveness of the method and the scientists who may distrust a technique which is becoming more and more esoteric.

POST-GRA... TRAINING COLLEGE
सत्यमेव जयते
Banipur, Po. Balgachi

Chapter I

THE LOGIC OF FACTOR ANALYSIS

Because factor analysis is a mathematical technique it is often thought that the non-mathematician is quite unable to assess its inner logic. Strangely enough it is often the same people who are most inclined to belittle the value of the method. It would almost appear that such people are under the impression that mathematics is in some way substituted for logical reasoning, and, like most substitutes, must be inferior. Psychologists, in particular, are apt to feel that this involves an unjustifiable invasion of the psychological field by mathematicians who do not understand the subject.

In actual fact the factorial technique is based upon simple logic and may be discussed in simple logical terms. It is simply a means of applying certain logical tests to somewhat complicated data and it is certainly no substitute for a thorough knowledge of the subject matter of the science concerned. It provides no more excuse for mathematicians taking over sciences such as psychology and biology than the use of words by historians constitutes a right of the linguists to take over history.

What then is the logical basis of factor analysis? Our simplest answer will, perhaps, be to state a number of simple axioms which will make clear the underlying reasoning:

1. Two tests which measure the same thing must give similar results. If we measure a table with a centimetre rule and then with a foot rule we shall get equivalent results. If we measure ten tables with each rule and place them in order of size we shall get the same order from both rules. If, however, we should measure height with one rule and length with the other we should not necessarily get this correspondence since we are measuring something different in each case.

2. Two tests which involve some common abilities will give results which agree to the extent of the common abilities involved. We may thus expect that the extent of the agreement will show the extent to which we are measuring the same things. This is the sort of result we should get if we compared lengths of table tops with perimeters of the same table tops. Since length is the most important element in perimeter there would be considerable correspondence.

3. Now suppose that test A and test B show some extent of agreement in this way (we need not worry yet as to how this is assessed). Suppose further that test A and test C show some agreement in this way too. We have grounds for suspecting that the three tests are measuring the same thing, that we have a common factor at work. We should, if this is true, find some agreement between tests B and C. If we do not find this we may conclude that B and C agree with different parts of A. If B is a test of memory and C is a test of intelligence this state of affairs could be easily explained by assuming that A is a test of both intelligence and memory.

Such reasoning as this differs in no way from what we should use in any experimental investigation. The only mathematical function involved is that of calculating the correlation coefficient to indicate the degree of relationship. The rest is just a matter of logical deduction from the linkages thus demonstrated. The techniques of factor analysis employ no magic. Their function is to show in quantitative terms the pattern of linkages among our variables. In this way insight is facilitated.

Students who are not adept in mathematics are inclined to be awed by the statistical juggling of the factorist, but if they bear in mind that the principles which we have just considered here are basic to it all they should be encouraged to try to master the arithmetical methods which we employ to deal with the various practical difficulties. In the pages which follow we try to explain these in a way which we hope will be comprehensible to anyone who has learned to carry out addition, subtraction, multiplication and division of decimals. Most of the examples have been worked to only two decimal places in order to keep them to a simple form.[1] It is hoped that those who have mastered the operations of multiplication and division with a slide rule will be tempted to follow the working. This use of the slide rule can be learned in a few minutes and anyone who is interested in trying factor analysis on his own account is strongly urged to get a rule and use it from the first.

[1] In practical work on psychological material it is seldom necessary to work to more than three decimal places and two places will often suffice, especially when group methods are employed.

Chapter II

CORRELATION

The first step towards a technique of factor analysis was made when Francis Galton, probably the greatest genius in the history of English psychology, introduced the function which is now known as the coefficient of correlation. The measurement of correlation is basic to our method. It provides us with an arithmetical statement of the extent to which two sets of measurements agree.

If we weigh a number of articles on a chemical balance and then repeat our weighing we expect to get the same results the second time. It is possible, however, that some of our results might differ by a small fraction. If instead of weighing the articles we asked someone to guess the weights it is likely that two series of guesses would show some very big discrepancies. All the method of correlation does is to provide us with a convenient scale for the measuring of such discrepancies.

Obviously there are many ways in which we might measure this sort of thing. We might, for example, simply find the difference between each pair of estimates and add all these together. It will be evident, however, that if we do this with regard to a series of weights and then again with measurements of reading ability the two sums will not be at all comparable. It is essential that our coefficient of correlation should have the same significance no matter from what sort of material it is derived. What is a high figure for twenty weights must also be a high figure for a thousand intelligence test scores.

The method which has been found to give the most effective results is known as the 'product–moment' method and is directly derived from the early work of Galton. It provides us with a measure which varies from $+1{\cdot}0$ for perfect positive correlation to $-1{\cdot}0$ for perfect negative correlation, with complete lack of correlation indicated by zero. If we spend a little time in seeing how this correlation coefficient is derived we shall be in a better position to understand what is involved in our factor analysis, which works with similar concepts.

PERFECT POSITIVE CORRELATION

Consider the two series of test results below in which scores vary from 1 to 5:

X: 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5.
Y: 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5.

We have here a perfect positive correlation. For every score on X there is a corresponding Y score which is exactly the same. We can represent this position neatly in what is called a *scattergram* (Fig. 1):

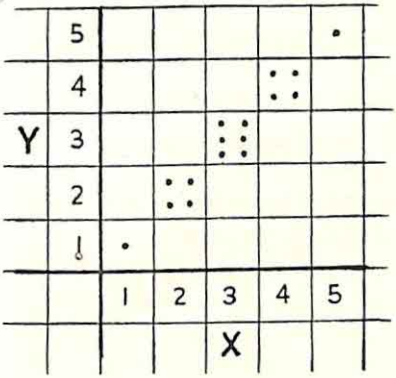

Fig. 1

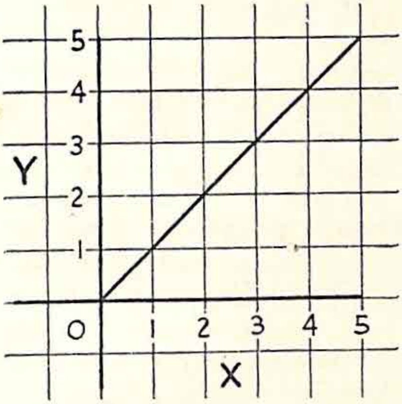

Fig. 2

This scattergram consists of five columns and five rows corresponding to the scores 1 to 5 for X and Y respectively. The number of dots in any column represents the number of times that score occurs in the X series. In the same way each dot in a row represents the number of times that the score occurs in the Y series. By this arrangement we are able to represent the score for the two series by the same dot.

It will be noticed that the dots arrange themselves in a diagonal pattern across the scattergram. This can be expressed quite well in a simple graph from which, for every X-score, we can read off the corresponding Y-score (Fig. 2).

PERFECT NEGATIVE CORRELATION

In order to get a perfect negative correlation we simply reverse our Y series. The scores would thus read:

X: 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5.
Y: 5, 4, 4, 4, 4, 3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 1.

Below we set this out in a scattergram and also represent it by a simple graph (Figs. 3 and 4). We have a similar diagonal arrangement to that of the positive correlation, but this time the diagonal runs from top left to bottom right.

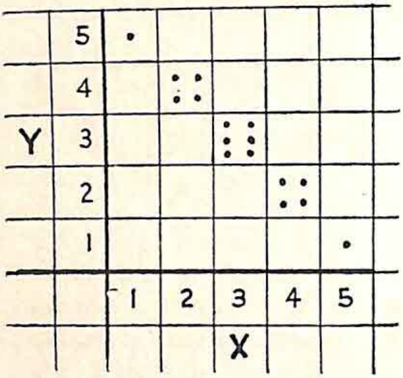

Fig. 3

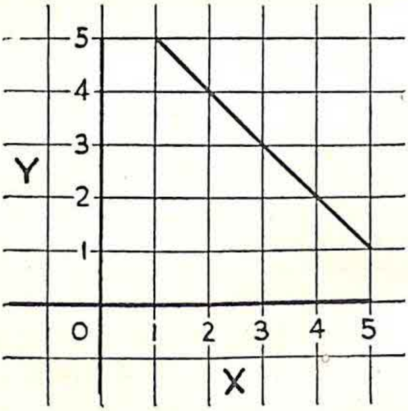

Fig. 4

ZERO CORRELATION

Let us now consider a case in which this neat sort of correspondence does not exist:

X: 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5.
Y: 3, 2, 3, 3, 4, 1, 2, 2, 4, 4, 5, 2, 3, 3, 4, 3.

As a scattergram this takes a form in which the diagonal arrangement is completely absent (Fig. 5).

How can we represent this in a graph? We must resort to the use of averages. Let us graph the average score for each column of the scattergram. That means that against each X value we graph the corresponding *average* Y *value*. The result is a neat horizontal line. If we do the same thing for rows, plotting *average* X *values* against each Y value we get a vertical line (Fig. 6).

These two lines are known as *regression lines*. It will be noticed that with perfect correlation, either positive or negative, we have only one

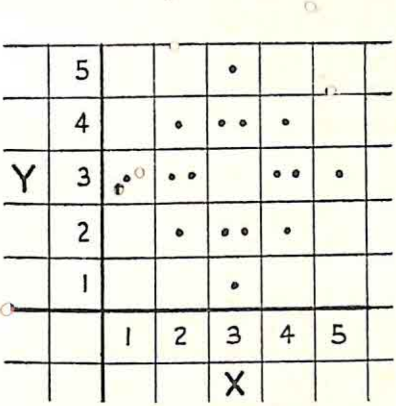

Fig. 5

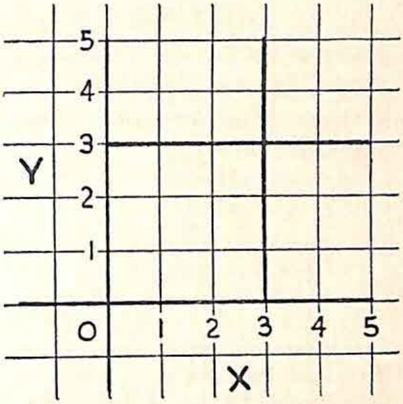

Fig. 6

regression line instead of two, or rather that the two regression lines coincide. Look again at the scattergrams for these two perfect correlations and it will be seen average X against Y is exactly the same as average Y against X.

It will be seen now that the regression lines can vary from the coincidence of perfect positive correlation (when they both run diagonally to the right and upwards) to the coincidence of perfect negative correlation (when they run diagonally from upper left to lower right). Midway between these two positions we have maximum divergence where the lines are at right angles. *This position represents zero correlation.*

The correlation coefficient is determined by the extent to which the two regression lines depart from the horizontal and vertical positions (i.e. the zero position).

Now from the diagonal position to the vertical position represents half a right angle or 45 degrees. The reader will see that it would be possible to express correlation as an angle which can take some value between zero and 45 degrees. It is more convenient, however, to use the tangent of the angle rather than the angle itself. The tangent of an angle of 45 degrees is unity, and the tangent of zero angle is zero. This is illustrated in the following series of diagrams. Let us place the angle R in a right-angled triangle and letter the sides a, b and c as shown.

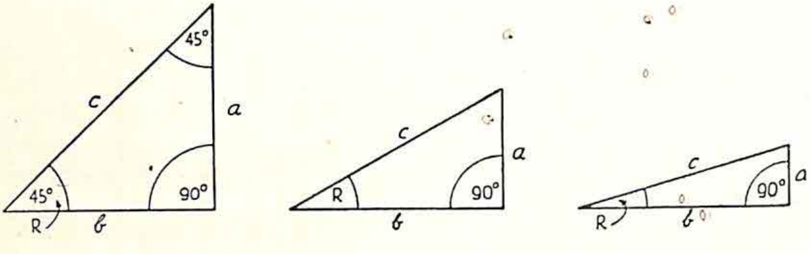

Fig. 7

Then the tangent of R is defined as the ratio $\frac{a}{b}$. If R=45 degrees, then a=b and the tangent of R will be unity. If R is less than 45 degrees, as in the second diagram, then a is less than b, and the tangent of R, that is, $\frac{a}{b}$ will be less than unity. The third diagram shows the angle R approaching zero, and the ratio $\frac{a}{b}$ has become smaller. Clearly when R vanishes, a becomes zero, so that the tangent of zero angle is zero. Instead of using an angular scale of 0 to 45 degrees we could use a tangent scale of 0 to 1·0, and this is just the sort of scale we need. Our correlation coefficient is thus expressed as a decimal fraction, negative or positive, which shows the approach to perfect agreement as it becomes nearer unity.

Since there are two regression lines we shall have two tangent measures of departure from the zero position, one measured from the vertical and one from the horizontal. These two measures are usually somewhat different, so again we have to resort to averaging. We take the mean of the two tangents, but not the arithmetical mean. Instead we use the geometrical mean. Instead of adding the two results together and dividing by two we multiply together and take the square root.

Intermediate Positive Correlation

Let us now consider the sort of case we are likely to meet in practice:

X: 2, 2, 5, 3, 2, 4, 3, 3, 3, 3, 1, 4, 3, 3, 3, 4.
Y: 1, 3, 4, 4, 3, 3, 3, 4, 3, 2, 2, 3, 2, 4, 2, 5.

Below we set this out in scattergram form and plot its regression lines.

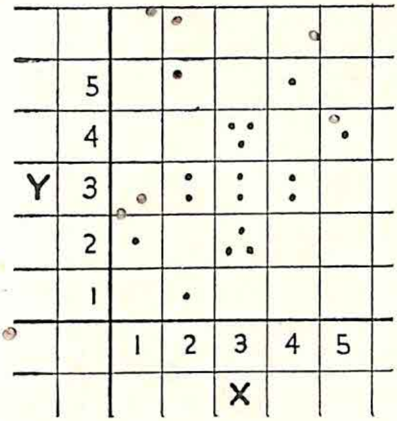

Fig. 8

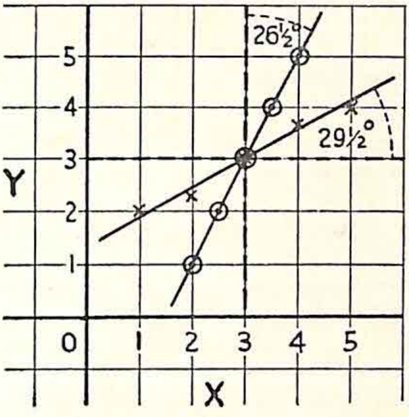

Fig. 9

It will be noted that on our graph we have drawn, through the point where the regression lines intersect, a vertical and a horizontal line so that we can measure more conveniently the two regression angles. These are 26½ degrees and 29½ degrees approximately and these correspond to tangents of ·50 and ·57 whose geometrical mean is ·53 (approx.).

ARITHMETICAL CALCULATION

This may seem a rather clumsy way to work out a correlation coefficient and it is. We have done it this way only because this best shows the significance of what we are doing, and it is the significance rather than the method that we are interested in at this stage. To get exact results we can work more exactly by an arithmetical method. In this way we can avoid two major difficulties. The first arises from the liability of graphical measurements to small errors and the other from the fact that when we plot the mean of our rows (or columns) we do not get a number of points in a straight line, but have to draw our line so that it comes as close as possible to as many points as possible. Theoretically this line should be so drawn that the sum of the deviations of the points should be at a minimum. The deviations are first squared which has the effect of making them all positive. But of course we cannot waste our time trying out various positions for the regression line, squaring the deviations and then adding them up to see which position gives the smallest sum. Instead we set about calculating the tangents of the regression lines directly.

1. We find the means of the two series.

2. We subtract from each score the mean of its series. The two resulting series of deviations from means we denote by a small x and y.

3. We now multiply together the corresponding x and y deviations, pair by pair and add the products.[1]

4. We square all the x and y deviations and sum each series.

The two tangents are now $\frac{\text{sum xy}}{\text{sum x}^2}$ and $\frac{\text{sum xy}}{\text{sum y}^2}$.

We have now only to note how the formula for the correlation coefficient is obtained from the two regression formulas:

r (symbol for correlation)=geometric mean of regression coefficient

$$=\sqrt{\frac{\text{sum xy}}{\text{sum x}^2}\times\frac{\text{sum xy}}{\text{sum y}^2}}=\frac{\text{sum xy}}{\sqrt{\text{sum x}^2\times\text{sum y}^2}}$$

We set out below the calculations involved in arriving at the two regression coefficients and the correlation coefficient. Of course there are no awkward decimals to be handled here. The figures are small and the means come out in whole numbers. But there are tricks for dealing with difficult material and we shall mention something of these shortly:

X	Y	x	y	xy	x^2	y^2
1	2	−2	−1	2	4	1
3	2	0	−1	0	0	1
4	5	1	2	2	1	4
4	3	1	0	0	1	0
3	2	0	−1	0	0	1
3	4	0	1	0	0	1
3	3	0	0	0	0	0
4	3	1	0	0	1	0
3	3	0	0	0	0	0
2	3	−1	0	0	1	0
2	1	−1	−2	2	1	4
3	4	0	1	0	0	1
2	3	−1	0	0	1	0
4	3	1	0	0	1	0
2	3	−1	0	0	1	0
5	4	2	1	2	4	1
Totals		0	0	8	16	14

---

[1] The meaning of 'product-moment' used in the name for this method will now be understood. The 'moments' are the deviations about the mean and it is the products of pairs of these which we use in our calculations.

$$\text{Regression coefficient for X} = \frac{8}{16} = \cdot 50$$

$$\text{Regression coefficient for Y} = \frac{8}{14} = \cdot 57$$

$$r = \sqrt{\cdot 50 \times \cdot 57} = \cdot 53$$

or

$$\frac{8}{\sqrt{(16 \times 14)}} = \cdot 53$$

PRACTICAL METHODS

In practice we seldom find a mean which comes out to a whole number but we need not worry about this. We can use the nearest whole number and make appropriate corrections for our totals. If we have sixteen figures in our series as above, our error will have been repeated sixteen times. If the whole number used for the arbitrary mean for the x column were ·5 too high this would mean a total error for the sum of x^2 of 16 times the square of ·5. This we should subtract from the x^2 column total. If the Y column required an arbitrary mean to avoid the use of decimals we should make a similar correction. A corresponding correction must be made for the sum of xy. Here, instead of squaring our errors we multiply the X error by the Y error and multiply this product by sixteen. We must be very careful here about the sign of our error. We subtract the arbitrary mean from the true mean, never vice versa, and if the answer is negative the xy correction works in the opposite direction. The sign has no effect on the x^2 and y^2 corrections since the square of a negative number is positive.

All this is summed up in the somewhat imposing formula below. We write x′ to denote deviations from the arbitrary mean to distinguish from deviations from the true mean which we represented by x. In this way we remember that the correction must be made. The errors involved in using the arbitrary mean we denote by $\bar{x}'$ and $\bar{y}'$. N is the number of measures involved.

$$r = \frac{\text{sum } x'y' - N\bar{x}'\bar{y}'}{\sqrt{(\text{sum } x'^2 - N\bar{x}'^2)(\text{sum } y'^2 - N\bar{y}'^2)}}$$

The correlation coefficient worked out in this way is the most reliable one we can use for most purposes but it involves a considerable amount of labour if our figures are large. If we have to calculate a considerable number of such correlations, as is usually the case in factor analysis, the method may not be practicable without the use of elaborate calculating devices. Increasing use has therefore been made of a much simpler method.

TETRACHORIC CORRELATION

The easiest form of correlation to use in practice is one which involves simple arithmetic and then the reading of the answer from a graph. In order to apply this tetrachoric method one divides each series of measures into high and low and sets them out in a table as below:

		Y		
		High	Low	Total
X	High	12	8	20
	Low	18	12	30
	Total	30	20	50

If we have a perfect correlation this table will be in diagonal form again. If the correlation is a perfect positive one we shall have entries only in the high-high and low-low positions while for perfect negative correlation all the entries will be in the high-low and low-high positions. With zero correlation there will be equal entries in all the squares.

Various methods have been devised for deriving a correlation coefficient from such a table as this. Usually they involve converting three proportions from the table into percentage form and then reading the result from a graph. If one takes care to have equal numbers of highs and lows in each series, which usually means taking the average as the dividing point, one can reduce the number of percentages to be calculated to two. In this way a large number of correlations can be handled without undue labour. Furthermore, it is possible by this method to correlate the yes-no answers to items in a questionnaire. One just considers the yes-answers as corresponding to high and the no-answers as low.

Chapter III

FACTORS

CORRELATION BASIS

The basis of factor analysis is that if two activities involve a common element there will be a correlation between them. If we apply two psychological tests to a number of people the correlation between the results will depend upon the extent to which the tests are calling upon common abilities.

Let us make this clearer by taking a physical ability. Suppose we have a number of taps of different sizes all attached to a water supply of constant pressure. We now apply two tests to these taps. The first test is the time required to fill a jug with water. The second test is the time required to fill a bucket with water. It is obvious that the taps which fill the jug quickly will also fill the bucket quickly and vice versa. If we arrange the taps in order according to the time required to fill the jug and do the same with regard to the time required to fill the bucket the two rank orders will be exactly the same. Our correlation will be perfect, 1·0.

But now let us introduce some complications. Suppose there is a choppy wind blowing while we are carrying out our tests. Not all the water goes into the receptacles. We shall not expect to find a perfect correlation now because the wind cannot be depended on. If its effect were constant we should have just a reduction of the amount of water flowing into the receptacles which would be equivalent to a reduced pressure. But our wind is very erratic. It may hardly blow at all during one filling and come with a tremendous gust during the next. On the whole, however, it blows away proportionately more water when we are filling the jug (because it is smaller) than when we are filling the bucket. Let us suppose that the interference with the bucket filling amounts to 30 per cent and the interference with the jug filling amounts to 50 per cent. Under these conditions tap size determines the bucket-filling speed to the extent of 70 per cent and the jug-filling speed to the extent of 50 per cent. Our correlation will now be 70 per cent of 50 per cent or ·7 of ·5 which gives us ·35. This is the correlation due to the factor of tap size and the extent to which the factor is operating with regard to the two tests is called the factor loading (·7 for bucket and ·5 for jug). The correlation is the product of these two loadings.

It will be seen that we have here introduced a very important new principle. The correlation can only be between two tests; the factor loading tells us something about *one* test. It gives us information which can be used to predict the results of a number of possible combinations.

All this will become clearer with a further example. This time we shall take a number of boys as the subjects for our tests. For our first test we ask them to empty buckets of water by using spoons. Some of the boys are considerably faster than others. We set a similar task for our second test but this time the buckets are filled with oil. We find that the results of the second test are not quite the same as those for the first. The speed factor does not show up quite so well with oil as with water. It is even less effective when we use sand for our test and when we come to treacle there is a further drop in efficiency.

Let us suppose that our speed factor operates to the extent of ·9 with the water, ·8 with the oil, ·7 with the sand and ·6 with the treacle. Taking these as factor loadings we get the correlations set out in the table below:

		water ·9	oil ·8	sand ·7	treacle ·6
water	·9	x	·72	·63	·54
oil	·8	·72	x	·56	·48
sand	·7	·63	·56	x	·42
treacle	·6	·54	·48	·42	x

The figures set out in the body of the table are the correlations which we should get from the factor loadings set out above and at the side. The crosses running diagonally occupy the position where we should find the correlation of the test with itself. We make no entries here for the time being.

If we examine this table carefully we find a very interesting relationship present. Consider the four correlation coefficients in the right top corner:

$$\begin{array}{cc} \cdot 63 & \cdot 54 \\ \cdot 56 & \cdot 48 \end{array}$$

If we multiply together the diagonally opposite coefficients we get:

$$\cdot 63 \times \cdot 48 = \cdot 3024$$
$$\cdot 56 \times \cdot 54 = \cdot 3024$$

Spearman, who first called our attention to this sort of situation, calls four such correlations a 'tetrad' and the difference between the two products the 'tetrad difference', which in this case is zero.

If we take any other four symmetrically arranged correlations from this table we will find that they also have zero tetrad differences, e.g. $\cdot72 \times \cdot42 = \cdot63 \times \cdot48$. That this must be so is easily proved by considering how the correlations are obtained from the factor loadings. If we substitute for the correlations the factor loadings from which they are derived we get:

$$(\cdot9 \times \cdot8) \times (\cdot6 \times \cdot7) = (\cdot9 \times \cdot7) \times (\cdot6 \times \cdot8)$$

Re-sorting and removing the brackets gives us:

$$\cdot9 \times \cdot8 \times \cdot7 \times \cdot6 = \cdot9 \times \cdot8 \times \cdot7 \times \cdot6$$

which obviously must be true.

Bearing this tetrad situation in mind let us now consider the diagonal entries which we left empty. It becomes obvious at once that if we did not know the factor loadings for a table of empirical correlation coefficients we should here have a means of determining them. The correlation of the test with itself will be the square of its factor loading. If, then, we can find the self-correlation we have simply to take its square root and we have the factor loading. This is easy. We just select a tetrad which includes a diagonal cell:

$$\begin{array}{cc} \cdot56 & \mathrm{x} \\ \cdot48 & \cdot42 \end{array}$$

Here we know that the missing number, when multiplied by $\cdot48$ must be equal to $\cdot56 \times \cdot42$, i.e. $\dfrac{\cdot56 \times \cdot42}{\cdot48} = \cdot49$. The root of this is $\cdot7$ which is the factor loading for the sand test.

In this same way we could calculate the factor loadings for the other tests so it would appear that, given empirical correlations, we could easily calculate the factor loadings. There is a catch to it, though. Each diagonal cell can be included in a number of tetrads (depending on the size of the table) and no matter which tetrad we use as the basis of our calculation we should get the same result. With the present table this can easily be proved to be the case. But this table was built up theoretically. With actual correlations the position is very different. Each tetrad may give us a different result. Which are we to take as correct? None of them

6656
7.1.94

will be quite correct but if we take the average of them all we shall be as near as possible to the right answer. Spearman's method of analysis was just a quick method of getting this average answer. We shall not consider it here because it is now never used. Instead we shall explain an equivalent and somewhat simpler method which was developed in England by Burt as the 'method of simple summation' and in America by Thurstone as the 'centroid method'.

Before going further, however, we should remark that Spearman used his method to deal with only one factor and used the tetrad criterion to prove that only one factor was present in a set of correlations. We shall have more to say of this when we have considered what happens with several factors present.

CALCULATING FACTOR LOADINGS

We now consider the centroid method of calculating factor loadings. This will be easier to understand if we set out a correlation table with letters to represent the factor loadings and combinations of these letters to represent the correlations and totals. This will make the relationships of the various parts of the table evident at a glance:

Loadings:	A	B	C
A	A×A	A×B	A×C
B	B×A	B×B	B×C
C	C×A	C×B	C×C
Totals: A+B+C	(A+B+C)×A	(A+B+C)×B	(A+B+C)×C

It is obvious at once that the correlation coefficients in any column are all multiples of the factor loading for the column and that the *column sum is equivalent to the sum of all the loadings multiplied by the factor loading for the column.*

It follows that, if we can find the sum of the factor loadings, it will be an easy matter to find the individual loading for each column. All that we need do is divide each column total by this sum of loadings. How can we find this sum of loadings? It is quite simple. Our grand total will consist of

$$\begin{array}{rl} & (A+B+C) \times A \\ \text{plus} & (A+B+C) \times B \\ \text{plus} & (A+B+C) \times C \\ \hline \text{Total:} & (A+B+C) \times (A+B+C) \end{array}$$

This grand total is thus the square of the divisor we are looking for so we can just take its square root and divide into the column totals. Let us try it out on a simple example:

	·25	·20	·15	
	·20	·16	·12	
	·15	·12	·09	
Sums:	·60	·48	·36	Grand total$=1\cdot44=1\cdot2^2$
Dividing by 1·2,	·5	·4	·3	

By using these factor loadings we find that we reproduce the correlation coefficients exactly.

All this works out perfectly but the reader may have noticed that we have cheated here. What about the diagonal entries? In practice we should not have them but we have used them here. Where did they come from? We must confess that we just calculated them theoretically like the rest of the table. This method, then, cannot be used for practical material? As a matter of fact it can. We just have to make a good guess as to what ought to go in these cells. With a large table a slight error here will not matter very much. With small tables we can gain more precision by repeating our work using the square of the loadings found in place of our guessed entries. By repeating this often enough we may calculate our results with any desired degree of precision.

With regard to guessing the 'communalities' as the diagonal entries are called (because they are measures of what each test has in common with the other tests) we should note that Thurstone has simply made a practice of choosing the highest correlation in the column and using this as a first approximation. For the largest column this is usually too low and for the smallest column it is usually too high, but if the loadings are calculated a second time, which does not take long, it is simplest to put in the highest coefficient as the communality.

More Than One Factor

Now let us consider a second factor. Suppose that the matrix we have just analysed was produced by giving a number of girls three tests, consisting of emptying basins of milk, cream, and syrup respectively. For the girls doing these tests, the speed factor produces loadings of ·5, ·4 and ·3. Now suppose that we give our girls spoons of different sizes. Their results will now depend not simply upon their speed but upon the size of spoon too. For the sake of illustration we shall assign loadings to the spoon factor and so build up another theoretical table of correlations. We get a table as below:

	·3	·4	·5
·3	·09	·12	·15
·4	·12	·16	·20
·5	·15	·20	·25

We add the correlation coefficients in this table to those of the table we have just analysed to get the amount of correlation which will result from the two factors. We get the table below which we proceed to analyse as before:

	·34	·32	·30	
	·32	·32	·32	Two factor table[1]
	·30	·32	·34	
S:	·96	·96	·96	$SS=2{\cdot}88$ $\sqrt{2{\cdot}88}=1{\cdot}697$
Loadings:	·56	·56	·56	

The next step is to use these factor loadings to calculate the correlations they will account for:

	·56	·56	·56	
·56	·32	·32	·32	
·56	·32	·32	·32	First factor matrix[2]
·56	·32	·32	·32	

This we subtract from the table above to get:

·02	·00	—·02	
·00	·00	·00	First residual matrix
—·02	·00	·02	

Our residuals here are so small that in practice we should ignore them as due to chance (the wind was blowing!). We should thus stop with only one factor. In this case we know there is no chance factor involved so we shall proceed and see what happens.

---

[1] Using S to represent Sums and SS the Grand Total.

[2] The inquisitive reader who squares ·56 will get ·31 as the best approximate answer to the second place. In fact, however, ·56 is itself an approximation and ·32 is the best approximation to the square, which will normally work to the nearest second place in this book.

First of all we notice that some of our coefficients are negative and when we add up our columns the totals are all zero. Now what do we do? We can get over this but before proceeding we must observe that these results show that we have not made any mistake in our arithmetic. If the columns of our residual matrix do not add up to zero (apart from rounding errors) we have made some mistake.

In order to proceed we must get rid of some negative entries. We do this by changing the sign of one or more tests. There is no reason in the world why we should not reverse the results of an intelligence test and call it a test of stupidity! The only difference will be that the people who were top now go to the bottom and tests with which this one had a positive correlation will now have a negative one.

Thurstone chooses the column which has the highest total regardless of sign and changes all the negative signs in this column to positive. To keep things even, whenever he changes a sign in this way, he has to change all the other signs in the same row and column.

In our present example we have two columns with similar totals when added regardless of sign. We choose the third where we find that there is a negative in the first row. To get rid of this we change all the signs in the first row and *first column*. This makes all our coefficients positive but we must remember that our first test is now a negative one. Note that the diagonal cell changed from positive to negative when we changed the row but changed back to positive when we changed the column. In practice we should not have touched this cell because in the first place the communality would not be entered (we guess this afresh for every factor) and in the second place it must be positive since it represents the *square* of the factor loading. We leave it in for this example because, since we are dealing with a theoretical table where everything must work out exactly, it happens to be the right communality. Now let us calculate our factor loadings for this second factor:

Sum columns: ·04 ·00 ·04; $SS=\cdot 08$; $\sqrt{\cdot 08}=\cdot 2828$
Loadings: ·14 ·00 ·14

Remembering that our first test is negative we now change it back and make the factor loading negative instead. Now let us set out the factor loadings we have obtained and compare them with the loadings from which we built up our two factor table:

Factor Loadings

Original		Found	
I	*II*	I	II
·5	·3	·56	—·14
·4	·4	·56	·00
·3	·5	·56	—·14

Obviously the process is a complete failure now that we have two factors involved. There is no correspondence between the two sets of factor loadings. But let us not give up hope yet. Often we can learn more from a graphical representation. Let us try this with our two sets of factor loadings. We measure factor I vertically and factor II horizontally:

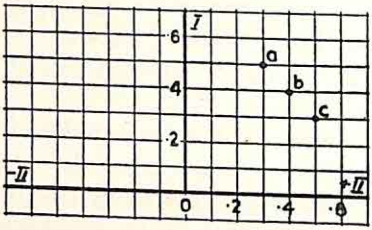

Fig. 10.—Graph of Original Factor Loadings

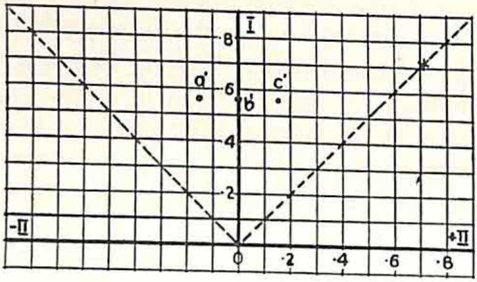

Fig. 11.—Graph of Factor Loadings from Analysis

An examination of these two graphs will show a very interesting situation: by rotating the axes of the second graph through an angle of 45 degrees it becomes identical with the other. The reader can easily verify this by copying the two graphs, superimposing the one on the other, placing a pin through the two intersections of axes and then, holding both papers up to the light, rotating one round the pin until the dots representing the factor loadings cover the shadow of the similar dots of the other.

We can convert the one set of loadings into the other by a simple arithmetical process. The change involved is one in the position of the factor axes so we just note the position of the new axes in terms of the old and adjust our loadings accordingly:

I rotated. We draw in the new reference axis at the required 45 degrees and mark a point on it one unit from the origin. In our diagram this is marked by a cross and this point, *read in terms of the original axes*, is ·71 along the I axis and ·71 along the II axis, so we multiply all our factor I loadings by ·71 and all our factor II loadings by ·71 and add the pairs together:

I × ·71:	·398	·398	·398
II × ·71:	—·099	·000	·099
Adding	·299	·398	·497

These loadings, apart from a small error due to the reading of our graph correspond to the original loadings on one of our factors.

We can get the results more accurately by calculating the new factor references instead of reading, but since the graph is our only guide as to the best rotation there is nothing to be gained. The reading is usually done in a slightly different way from what we have described, however, and this is sometimes more accurate. It will be as well to explain it here because we shall be using a similar principle in connexion with other methods of rotation.

Instead of measuring one unit along the new axis we mark the point directly opposite the unit position on the old axis and note the position of this point in terms of the other old axis. The diagram below makes this clear:

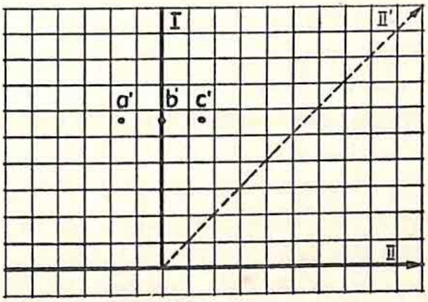

Fig. 12

The arrow head on axis II shows unit length; the arrow head on new axis II′ marks the point immediately above the unit position on axis II. The height of this arrow head is its position in terms of factor I. It is here exactly one unit itself. This tells us that the multipliers for loadings I and II should be in the proportion 1:1. The sum of the squares of the multipliers must be unity again, though, so we have to make an adjustment known as 'normalizing'. This is done by squaring the numbers, adding them together, finding the square root of their sum and dividing each by this:

	Squares	Divided by root
	1·0	·7071
	1·0	·7071
Sum:	2·0	
Root:	1·414	

By using the multipliers thus we get a nearer approach to our original loadings.

II rotated. We carry out just the same process for the new position for this reference axis. Our unit distance is taken along axis I. The corresponding distance along axis II is, however, in a negative direction this time: —1·0 instead of plus. So our multipliers become:

I	·7071
II	—·7071

So our factor loadings become:

I × ·7071:	·40	·40	·40
II × ·7071:	—·10	·00	·10
Subtracting:	·5	·4	·3

This gives us back the other factor exactly.

Note, however, that the rotated factors are in opposite order from the originals. The reason for this switch-over will be apparent later.

Chapter IV

THE PRINCIPLE OF ROTATION

We have demonstrated now that it is possible to take a table of correlation coefficients and, by carrying out certain mathematical operations, get back the factor loadings from which the correlations were derived. But it has to be noted that we were able to do this only because we knew what these original factor loadings were. Could we have succeeded without this knowledge? For this particular case the answer must be no, since there is nothing unique about the factor loadings with which we started. We just happened to choose these loadings and they have no superiority over any other set of loadings that would produce the same correlation coefficients. And be it noted that there is an infinite number of such possible loadings. For every position in which we like to place the reference axes we get a new set of loadings. That means 360 different positions if we vary by only one degree at a time and obviously we may vary by minute fractions of a degree.

We thus have, mathematically, an infinite number of solutions to our factor problem but psychologically (and the same applies to other sciences) we hope to find a unique solution. We expect our factors to correspond to psychological units, to indicate to us the fundamental dimensions of personality. Are there any criteria we can use to indicate whether or not we have rotated our factor axes to a position which will make the factors correspond to something real in this sense? We think we may claim that there are such criteria.

These criteria are simply the criteria which we use in all scientific explanation. Basically they are two:

1. The explanation should be as simple as possible.
2. It must be consistent with other explanations which we accept.

Obviously we cannot accept a number of independent explanations which contradict one another. Each individual explanation must be consistent with all that we know but within these limits it must be as simple as possible.

Newton's theory of gravitation supplied both a simpler explanation of many phenomena and one which was consistent with all known facts. When facts were discovered which did not square with the theory it had to be modified. Precisely the same applies to our factor approach.

We demand simplicity but consistency. Not all rotations are equally acceptable. We seek the rotation which will give us the simplest account of the material concerned but also one which can be squared with all that we know of the science concerned. This means that the introduction of factor analysis certainly does not mean handing over psychology or any other science to the mathematicians. It means more use of mathematical tools but progress will still depend upon a thorough insight into the science concerned. The use of a calculating machine does not change any fundamental principles!

SIMPLE STRUCTURE

But what do we mean by simplicity with regard to factor loadings? We owe most of our knowledge here to Thurstone who has developed a theory of what he calls 'simple structure' in the factor matrix. Whenever we carry out a centroid analysis we get one or more factors *which usually have loadings on every test involved*. Furthermore the first factor may have all positive loadings but all the other factors have both positive and negative loadings in about equal proportions.

Why should this happen? It is a necessary result of the mathematical procedure involved. Our first factor accounts for as much variance as possible by one factor. It gives us loadings which are bigger than justified by some of the correlation coefficients but smaller than would be indicated by other coefficients. It gives us the loadings which *on the average* are justified by the correlations. This means that about half of our residuals will be negative and so we get negative loadings for the second factor and the same happens with all succeeding factors.

But this sort of result is, of course, quite artificial and the result of our method rather than of something unvarying in our material. What sort of results should we expect on commonsense grounds? In the first place we should expect few negative loadings with much of our material. If we are studying abilities we shall not expect to find many which are capable of helping with some tasks and hindering with others. Certainly it would be strange if, apart from one ability, all functioned both positively and negative to about equal extents.

In the second place it seems unlikely that, if we apply many tests, *all* the abilities involved should be important for *all the tests*. It seems much more likely (unless the tests are very similar) that some groups of abilities should be concerned with some of the tests and other groups with other tests. In other words some of our tests should have zero loadings on some of our factors. It is just such a state of affairs which Thurstone seeks in simple structure. He rotates the factor axes so that there will be as many zero loadings as possible and as few negative loadings as possible.

Let us consider now an actual application of this approach to empirical material. Below we set out the factor loadings which were obtained by a centroid analysis of the intercorrelations of five tests:

	I	II
1	·88	·39
2	·78	—·35
3	·95	—·14
4	·95	—·25
5	·91	·35

It will be seen that we have the usual arrangement of a first factor which accounts for very much of our variance and which is positive throughout, with a second bipolar factor whose loadings are much smaller. We plot the loadings for the two factors:

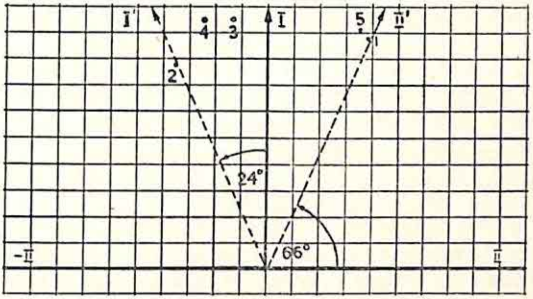

Fig. 13

It is evident that all the points representing the factor loadings fall within a narrow angle. We have only to rotate axis I anti-clockwise through an angle of 24 degrees to get rid of our negative loadings. At the same time we can rotate axis II through an angle of 66 degrees without introducing negative loadings. If we do this test 2 will fall on an axis and test 1 will fall on the other axis so that both will have one zero loading. Test 5 will also be very close to the new factor II axis and so will have a near zero loading. The two new positions are shown in the diagram by dotted lines and denoted I′ and II′. Note that the new axes are no longer at right angles. We shall speak of this later. In the meantime we must calculate the factor loadings for this rotated position.

Reference axis I rotated. We calculate the new factor loadings as already explained but with a slight modification. It will be remembered that in our first example the new factor loadings came out in reversed order. This would not happen here. Instead we should find the method had completely broken down. The trouble lies in a simple fact which is easily overlooked. When we measure from reference axes we use one axis as a base line and measure in terms of the other. At least this is the case when the axes are at right angles. When the axes are *not* at right angles we need two extra lines, one at right angles to each of the axes.

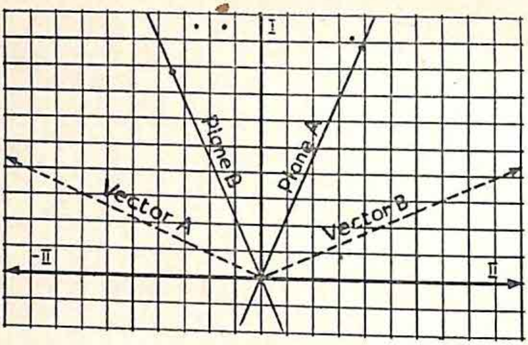

Fig. 14

These extra lines we shall call *planes* because they really represent planes at right angles to the reference axes. Now it is the position of the planes which we adjust with regard to the points representing the factor loadings on our diagram. With an orthogonal rotation the plane for one factor axis coincides with the other factor axis so we were able to ignore this complication provided that we made allowance for the reversal in the factor names. With oblique rotation it is necessary to distinguish between plane and reference axis. Figure 14 makes the position clear. The new reference axis is called a *vector* because it is of definite length, corresponding to unit length on original axis II. The letters A and B are used to denote reference to factor I and II respectively.

To find the new reference axis for factor I we rotate its position (originally coincident with axis II) to that in the diagram where it gives tests 1 and 5 zero or near-zero loadings. Its vector is at right angles to it, in the negative quadrant. The arrow tip on the vector gives us the coefficients for the new position: ·45 on I and —1·0 on II. We normalize these to make the sum of their squares equal to unity, as previously explained, and so we get the multipliers ·41 and —·91. Using these we get the new factor loadings:

I × ·41:	·36	·32	·39	·39	·37
II × ·91:	·36	—·32	—·13	—·23	·32
Subtracting:	·00	·64	·52	·62	·05

Factor II rotated. We do the same thing for our second factor loadings. Vector B gives us ·45 on I and 1·0 on II, the same figures as for factor I (because this rotation is symmetrical) but with the second coefficient positive instead of negative. Normalized these coefficients will be ·41 and ·91 and we again multiply the original loadings, but this time *add* the products:

I × ·41:	·36	·32	·39	·39	·37
II × ·91:	·36	—·32	—·13	—·23	·32
Adding:	·72	·00	·26	·16	·69

For convenience of reference we set out these factor loadings with their test numbers:

	I′	II′
1	·00	·72
2	·64	·00
3	·52	·26
4	·62	·16
5	·05	·69

It will be seen that we have three loadings which are equivalent to zero.

We have now to ask ourselves what these factor loadings mean. This we cannot do until we know the nature of the tests. They are not tests of the psychological variety but simply physical measurements. We took half a dozen sheets of notepaper of various sizes and made five different measurements of each:

1. Length
2. Breadth
3. Diagonal
4. Area
5. Length squared + breadth.

It is at once obvious that measurements of this kind can be reduced to the two fundamental ones of length and breadth. Our factors should correspond to these if we have rotated correctly. This means that they should be two in number (right), that the first test should have a high loading on one factor but not on the other (right), that the second test should be the reverse of this (again right), that the last test should have a high loading on the length factor and very little on the breadth factor since length squared would dwarf breadth (right again) and finally that tests 3 and 4 should have loadings on both factors (right but not correctly balanced). It appears, then, that our rotated solution does make good sense. Apart from length of diagonal and area, which seem to be too lightly loaded with the length factor the results are everything we could hope for. This is somewhat surprising under the circumstances since we took only six cases for our correlation coefficients, which makes them very unreliable, and we used a simple method of calculating which would lose all the finer distinctions.

It is time now to revert to the fact that the rotation which we carried out here makes the new factor axes depart from their right-angled relationship. They are now at an angle of 48 degrees instead of the original 90 degrees. What does this mean? It will be remembered that when we were discussing correlation we found that when the regression lines were at right angles there was no correlation and that, as the angle between them decreased, the amount of correlation increased until with zero angle (lines coinciding in one of the diagonal positions) the correlation became perfect, positive or negative. The same significance attaches to our factor axes. If they are at right angles there is no correlation between the two factors; they are quite independent. Our angle of 48 degrees here corresponds to a correlation coefficient of ·66.[1]

We have now to inquire as to whether we are justified in making a rotation which results in correlated factors in this way. The natural tendency is to look for uncorrelated (i.e. independent) factors but Thurstone points out that the dimensions which we are prepared to accept on other grounds are usually ones which are correlated. We are inclined to think that length and breadth are independent measures but in actual fact they are usually correlated as we have found here with regard to our sheets of paper.

---

[1] It should be noted, however, that the correlation between factors is not related to the angle between the reference axes in quite the same way as the correlation coefficient is related to the angle between regression lines. As will be seen later the actual correlation between factors has to be calculated in a different way.

Thurstone and Thompson have both measured boxes in a variety of ways and analysed the intercorrelations. Simple structure has in each case given three factors which could easily be recognized as length, breadth and depth but in each case the factors have been correlated. This is an extremely important fact in connection with the well-known controversy as to the existence of a general factor among mental abilities.

SPEARMAN AND g

Spearman was very dissatisfied with conflicting conceptions of intelligence and sought to give it a rigid definition. He noted that many kinds of mental tests showed significant correlation. Furthermore he found that when these intercorrelations were set out in a table they manifested the sort of hierarchical structure which we noted in our first one-factor table (p. 20). He concluded that there was a general factor present in all these tests and that this general factor was intelligence. So we have the famous 'two factor theory'. The amount of success achieved with any test was said to depend upon a general factor, which it shared with other tests, and a specific factor which was unique to the test. The specific factor did not produce any correlation since we have correlation only in so far as factors are common to two tests. Only when two tests were very similar was this likely to happen. It was recognized, however, that two similar tests could have common specifics. In fact this was made an excuse for eliminating certain tests when the Spearman theory was being tested.

It is interesting to notice some of the implications of this last statement. Similar tests would have unusually high correlation because of similar specific factors. They could not, therefore, be allowed in a battery of tests designed to test whether the two factor theory were true. But the test of similarity could only be that the tests *did* produce correlations not in accord with the theory! The position became somewhat clearer with the development of a technique for handling more than one factor. It then became obvious that as well as a general and a specific factor there were also group factors of varying degrees of generality to be accounted for. The status of the general factor, however, was sustained by centroid analysis since it invariably appeared under these conditions.

At this stage of the controversy we find two distinct schools of thought. The Spearman school, using a centroid or similar analysis, without rotation, finds a general factor invariably present. The Thurstone school, making use of rotation, arrives at factors which do not appear to have a general factor. What was needed here was not an evaluation of the results but an evaluation of the methods. Centroid analysis necessarily produced a general factor while oblique rotation (i.e. correlated factors) deliberately produced group factors.

The end of this controversy seems to be in sight as the result of the study of 'second order' factors. If factors are correlated it follows that we can analyse their correlation too and it is this which gives second order factors. In this way we find that after our rotation has abolished the general factor it comes back as a second order factor. With our analysis of notepaper we get a second order general factor which influences both length and breadth. In the boxes example we have a similar second order factor which is common to length, breadth and depth.

But if we get a general factor under these circumstances surely we have proved too much! We cannot expect anything equivalent to intelligence among our boxes! What can be the nature of a general factor here? The factors we should expect have already been accounted for. Is this simply an artifact? We doubt it. It would rather seem that it is a shape factor. With our paper we were not just measuring length and breadth: we were measuring rectangles, and rectangles involve more than length and breadth. So with our boxes. These measurements were not just haphazard ones but related to box shape. Boxes which are very long tend also to be rather wider and deeper than usual; deep boxes tend to be proportionately longer and wider; seldom do we get a box which has one measurement in fractions of an inch and another dimension expressed in yards. So there *is* a general factor for our boxes and papers and our analysis is justified.

But this makes us look at the general factor in the mental field in a new light. May it be that *g* is simply a general tendency for all mental dimensions to vary together, a sort of mental size factor? To some extent this may be so. We should be inclined to suggest that the Spearman *g* comes from three sources:

1. A general tendency for mental (and physiological) capacities to vary in the same way. This is equivalent to 'boxiness' in boxes. On the average we find that strong, healthy bodies go with high intelligence. Despite the popular belief, brains and beauty *do* go together. Such correlations are, however, rather small but they constitute the basis for a true general factor. But this factor has no significance at all from the point of view of intelligence. Spearman's notion of a mental energy which can function in a variety of machines (corresponding to special abilities) cannot, by any stretch of imagination be justified here.

2. An artifact due to the type of analysis used. It will be noticed that with the notepaper example our first factor was too big not merely because it included the size factor but because it included some of the variance which, after rotation, went to the second factor. We have already pointed out that a centroid analysis necessarily exaggerates the early factors at the expense of later ones in this way. We found the same thing to happen in our first theoretical example. Here so much was accounted for by the first factor that we hardly seemed justified in calculating a second at all. It will be found that, whenever a theoretical table of correlations is set up to represent two equal factors, a centroid analysis will give far greater weight to the first. It would seem, therefore, that a general factor obtained by the centroid method can be suspected of being exaggerated.

3. A wide group factor. We come here to the part of *g* which leads to its identification with intelligence. Of the precise nature of this we know relatively little yet. It is possible that we have here a group of intellectual abilities which are united by the group factor. It would seem feasible that factors should appear at different levels. There is evidence that our universe has something of a hierarchical nature. If this is so we can expect factors to reflect this and we should here have an explanation of the fact that we find first, second and even third order factors.

Chapter V

FINDING SIMPLE STRUCTURE[1]

We have already stressed the need to make sure that the factor reference axes are correctly orientated before attempting to interpret factor loadings. It *is* possible to interpret two and even three factor centroid solutions without rotation but this can be done only because one can foresee what would be the pattern after rotation and actual rotation can always be recommended as the safest procedure.

In particular there are two points that may be overlooked in a simple inspection of the centroid loadings of even two factors. In the first place, it is difficult to judge the relative size of the factors since the first factor contains both whatever general factor there is plus something from the group factors. Careful rotation is required to gain insight into this. In the second place, it is difficult to decide whether two centroid factors represent two overlapping group factors or two oblique group factors plus a general factor which accounts for the correlation between the two oblique factors.

Even in the case of a cluster-directed analysis[2] it is usually desirable to check by rotation before interpretation. The complexity of factor patterns is so great that it is important to make use of all possible criteria interpretation.

ROTATIONAL METHODS

We shall refer here only to the simpler methods of rotation. One we have already dealt with to some extent in discussing a two-factor rotation. If there are more than two factors the method can still be employed but the factors have to be taken two at a time and the process may have to be gone through many times to get precise results. Since an early error may be perpetuated right through the series it is preferable to use a method which makes each calculation on the original figures. The most effective of such methods is one that has recently been developed by Thurstone. It is simple to understand and requires little experience in application.

SINGLE PLANE METHOD OF ROTATION

This method concentrates on one factor at a time. Instead of making a rough rotation of all factors and then repeating the whole process it notes all the adjustments of a factor as suggested by its relationships to the other factors and combines these adjustments in a single operation. This process is repeated for the single factor until the rotation is sufficiently precise and then a second factor is taken. We need, therefore, explain the method with regard to one factor only. Other factors are handled in exactly the same way. We shall try it out on three factors obtained by the centroid method:

---

[1] This chapter can be omitted by readers who are not concerned to use factor analysis.

[2] See page 48.

Centroid factors	I	II	III
1	·68	—·24	—·47
2	·51	—·36	—·07
3	·58	·29	—·21
4	·54	·35	—·04
5	·36	·47	·26
6	·40	·31	—·00
7	·23	—·40	·27
8	·30	—·42	·26

Since these are centroid factors none is likely to correspond at all closely to its simple structure equivalent so we have first to select a trial position. This is conveniently based upon a test which has high loading on the centroid factor. We choose test 1 and use the three coefficients (factor loadings) for this test as our multipliers (having first normalized them) to find the new loadings for the factor. There are three multipliers because we have here three factors and we are finding simple structure with regard to the three factors at once. In multiplying here we take the positive or negative nature of our multiplier into consideration so that we simply have to add the columns. Previously we took our multipliers as positive or negative and then added or subtracted according to the sign. This made more evident what we were doing but in practice it is better to take note of the sign for each individual multiplication.

Factor A_1

I× ·78:	·53	·40	·45	·42	·28	·31	·18	·23
II×—·28:	·07	·10	—·08	—·10	—·13	—·09	·11	·12
III×—·54:	·25	·04	·11	·02	—·14	·00	—·15	—·14
Sums:	·85	·54	·48	·34	·01	·22	·14	·21

We now take these factor loadings and plot them against the three original factors. To save work we neglect all cases where the A_1 loading is not the smaller of the pair as these have little influence in fixing the reference axes:

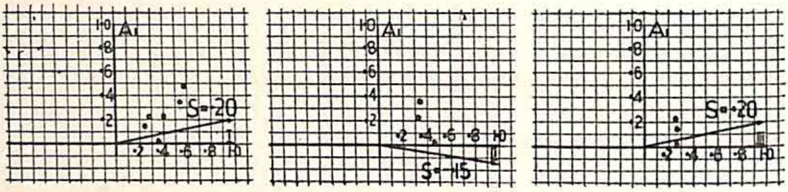

Fig. 15

We now draw in new planes for the A_1 factor on all our diagrams. It is not necessary to draw the vectors at right angles in this case as the subsequent calculation is based upon the slope of the new plane. This is read in just the same way as we took the height of the vector in the first method. We simply take the height of the point in the plane which is directly above the unit position in the base line. It is represented in the diagrams here by an arrow head again. For factor I it is ·20, for factor II −·15 and for factor III ·20. We now carry out the calculation shown below. The letters represent:

A_1 = normalized multipliers.
S = slope (i.e. height of plane above unit position on base.)
a = (A_1−S) divided by (1−A_1×S)
A_2 = a normalized (i.e. divided by root of sum of squares).

A_1	S	A_1−S	A_1×S	1−(A_1×S)	a	a^2	A_2
·78	·20	·58	·16	·84	·69	·476	·71
−·28	−·15	−·13	·04	·96	−·14	·020	−·14
−·54	·20	−·74	−·11	1·11	−·67	·449	−·69
						·945	= $\cdot 972^2$

We now use the A_2 multipliers to find the A_2 loadings just as we did for A_1. There is no need to set out this work in detail so we just give the new loadings:

Factor A_2: ·83 ·46 ·51 ·36 ·01 ·24 ·03 ·09

We plot these as before against our original factor loadings:

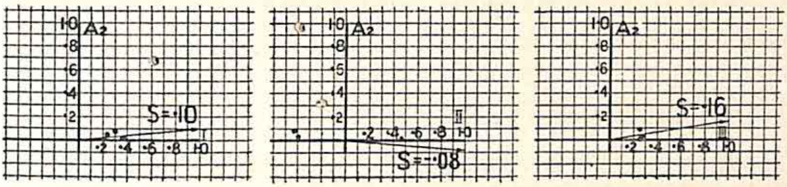

Fig. 16

It will be be noticed that we have omitted a few more points, those which fall well above the reference axis. These are not critical and it is a waste of time to plot them. The other points fall closer to our axis than before but a further adjustment seems desirable so we put in the new planes and read off their slopes. From these we calculate our new multipliers:

	A_2	S	A_2-S	$A_2 \times S$	$1-(A_2 \times S)$	a	a^2	A_3
I	·71	·10	·61	·07	·93	·66	·436	·65
II	−·14	−·08	−·06	·01	·99	·06	·004	−·06
III	−·69	·16	−·85	−·11	1·11	−·77	·593	−·76

$1\cdot033 = 1\cdot016^2$

And from these multipliers we get the new loadings:

Factor A_3: ·81 ·40 ·52 ·36 ·00 ·24 −·03 ·02

We plot these again to see whether they are precise enough:

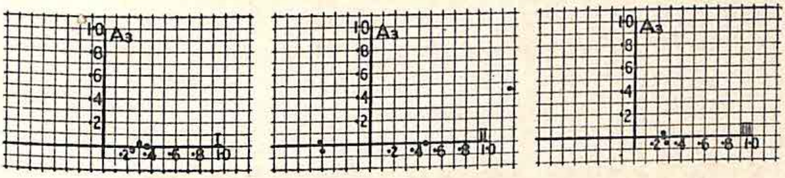

Fig. 17

It is obvious that our reference axes now go through the clusters of points and no further improvement is likely so we accept these loadings for our first factor. We should now proceed to rotate our other two factors in the same way. The reader may find it interesting to try this but we shall now go on to demonstrate another form of rotation.

METHOD OF EXTENDED VECTORS

This is by far the most convenient method to apply to the rotation of three factors since it finds the positions for the three planes in one operation and does not need repeating for finer adjustment. When applied to more than three factors it becomes more complicated and loses much of its superiority so we shall limit ourselves here to its three-factor application. Readers who wish to explore its possibilities further must consult more advanced texts.

The essence of the method is the projection of the three dimensions involved in three factors on to a flat surface. Two factors are represented by two lines on a flat surface. If the factors are orthogonal ones the lines are at right angles. A third orthogonal factor will be at right angles to the other two and so cannot be represented directly on paper since the third axis would need to project vertically from the paper.

We could make a model to represent the factor axes. The factor points would then be represented by little knobs on the ends of wires projecting from the centre. Suppose that we set up such a model and then enclose it in a large globe. Each wire representing a trio of factor loadings could now be extended to the surface of the globe and shown as a dot. We thus get a pattern on a curved surface and it is but one more step to convert this to a flat surface. Where all the factor loadings are positive or potentially so the factor points will fall within one quadrant and will easily project on to a plane surface.

This may seem to involve rather complicated operations but the practical steps are very simple. We reduce our three dimensions to two by the simple process of dividing one loading into the other two. This gives us the extended vectors which can now be plotted directly as before. The process will be clear from an actual example. We take the factor loadings of the example just considered and for each test divide the factor I loading into the factor II and III loadings and so get the extended factor matrix. For test 1 the factor II loading becomes —·24 divided by ·68 and the factor III loading becomes —·47 divided by ·68. For test 2 we divide by ·51 and so on. The result is:

Extended Factor Matrix

	II	III
1	—·35	—·69
2	—·71	—·14
3	·50	—·36
4	·65	—·07
5	1·30	·72
6	·77	—·00
7	—1·74	1·17
8	—1·40	·87

When plotted these take the form below (Fig. 18). Through these points we can now draw three planes instead of two and the planes no longer have to pass through the point at which the original axes cross. We may draw them wherever a number of points appear to be in line. If simple structure is present we should be able to draw a triangle with many of the points on the bounding lines. This works out quite nicely here. All the points near lines will have near-zeros for this factor; points falling at the corners of the triangle will have near-zeros on two factors and all points within the triangle will have positive loadings.

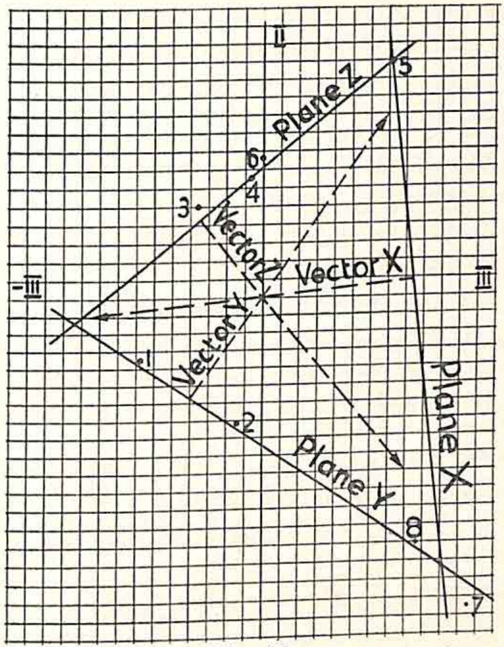

Fig. 18

The next step is to read off from the lines here drawn the appropriate multipliers to get the new factor loadings. We first draw a vector at right angles to each plane. This vector must pass through the point of intersection of the original axes and extend to a point above unit distance on the original factor axis to which it is most nearly parallel. We have put in these vectors. For plane X it lies −·09 units from the III axis. The III axis we take as unity and it is negative because the vector is pointing in a negative direction. The −·09 units are measured parallel to the original factor II axis so we set this down as the coefficient for II. The original factor I axis is to be regarded as extending vertically below the paper at the point of intersection. For each new reference axis, therefore, its factor I coefficient is simply its distance from this central point. This is measured at the point where the new axis cuts the old. For the X plane this is ·86. So the multipliers for the new X factor are these coefficients normalized as before: ·65 I, −·07 II, −·76 III. This is almost exactly the same as we obtained by the previous method so we may feel fairly confident that we have obtained the best rotation.

We go on to calculate the coefficients from which the multipliers for our other factors are obtained by the same process of normalizing. For convenience we tabulate the measurements:

Coefficient on I: distance from origin at which the plane crosses the axis nearest to parallel to its vector.

Coefficient on II: If this is the axis to which the vector is nearest parallel it is taken as unity, the sign depending on whether the vector is directed to a positive or negative quadrant. Otherwise the coefficient is read from the terminal point of the vector.

Coefficient on III: According to same rules as II.

We set out the three sets of coefficients in a table which, when normalized, gives us the transformation matrix (or sets of multipliers) which will convert our factor loadings to their simple structure form:

	New axes in terms of old			Transformation matrix		
	x	y	z	x	y	z
I	·86	·86	·70	·65	·58	·48
II	−·09	1·00	−1·00	−·07	·67	−·69
III	−1·00	·70	·80	−·76	·47	·55

The transformation matrix gives us three sets of multipliers. The three in the x column are similar to those already found by the single plane method so that, when we multiply as before, we get the same loadings for our rotated factor. We multiply by the three multipliers in the y column in the same way and sum to get the loadings for the second rotated factor. The same is then done with the z column. We thus get the new factor loadings below:[3]

Rotated Centroid Factors

	I	II	III
1	·81	·01	·24
2	·40	·02	·45
3	·52	·43	—·04
4	·36	·52	
5		·65	—·01
6	·25	·44	·02
7	—·03	—·01	·54
8	·02	·01	·57

It will be seen that we now have a minimum of three zero-equivalent loadings in each column and this suggests a meaningful pattern.

It is of interest to find to what extent our factors are now correlated. This we can do by finding the inverse of our transformation matrix, normalizing and then multiplying it by itself.

Inverse of transformation matrix. Unfortunately we cannot here use the simplified method that is possible with the cluster-directed method[4] because we do not have zeros above the diagonal. We proceed as follows:

1. Evaluate the first minors. To make this process clear we set out a matrix with letters to refer to the coefficients:

---

[3] This process is known as matrix multiplication and it may be useful if we explain here something of its general principles. Briefly each entry in the new matrix is the result of multiplying the corresponding *row* entries in the old matrix by the corresponding *column* entries in the multiplier and finding their sum. Thus the entry for row 1, column 1, is obtained by multiplying each item in row 1 of the original matrix by its corresponding item in column 1 of the multiplier. In our example here this gives three products which are then added. In practice it is more convenient to carry out all multiplications for a column by the one multiplier while the slide rule is set. This is what we have done in our example above. A little thought will show that it merely alters the order of the work. Students who are not familiar with matrix algebra will probably find it easier to think in terms of sets of multipliers, as we have referred to them in the text, each multiplier of the set being used in turn and the summation for the set completed before the next set is used.

[4] See page 61.

$$\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}$$

The first minor of a is the part left when we delete the row and column in which it occurs. This is:

$$\begin{array}{cc} e & f \\ h & i \end{array}$$

and to evaluate this we take (e×i)—(h×f).
The minor for b is:

$$\begin{array}{cc} d & f \\ g & i \end{array}$$

and we find the difference between the diagonals again as above. Great care must be given to the signs here. When the matrix is completed we reverse every second sign, counting from the end of one row to the beginning of the next.

Applying this method to our transformation matrix we get:

$$\begin{array}{rrr} {\cdot}69 & {\cdot}56 & {\cdot}46 \\ -{\cdot}10 & {\cdot}72 & -{\cdot}75 \\ -{\cdot}72 & {\cdot}42 & {\cdot}48 \end{array}$$

2. We should now divide each element by the value of the determinant but as we are going to normalize the columns of our matrix we can omit this step and proceed at once to normalize.

Inverse normalized. This is done, as before, by squaring, finding the square root of the sum of each column and dividing into every element in the column. This gives us:

$$\begin{array}{rrr} {\cdot}69 & {\cdot}56 & {\cdot}42 \\ -{\cdot}10 & {\cdot}72 & -{\cdot}74 \\ -{\cdot}65 & {\cdot}42 & {\cdot}48 \end{array}$$

Normalized inverse multiplied by itself: This can be done by multiplying each element in a column by the corresponding element in each column in turn, including its own.[5] So:

1. New diagonal elements will be the sum of the squares of the corresponding columns. This gives the correlation of I with I, II with

---

[5] It should be noted that this applies to this method of calculating an inverse but that when other methods are used it is the *rows* which should be multiplied together. See page 62 for an example where the transformation matrix has to be read in terms of rows because of a different method of calculating the inverse. Strictly, we should interchange rows and columns in the example above to get the true inverse. It would then come into line with other methods and be read in terms of rows like later examples.

II and III with III. It must, of course, be unity and that is why we normalized our columns.

2. The correlation of I with II comes from multiplying elements in column I with column II and adding them:

$$\begin{array}{r} \cdot 69\times\cdot 56 = \cdot 39 \\ -\cdot 10\times\cdot 72 = -\cdot 07 \\ -\cdot 72\times\cdot 42 = -\cdot 30 \\ \hline \cdot 02 \\ \hline \end{array}$$

3. Correlation of I with III and II with III is obtained in the same way:

In this way we get the table below:

Intercorrelations of Oblique Factors

	I	II	III
I	1·00	·02	·05
II		1·00	—·06
III			1·00

This is almost a negligible amount of correlation but it must be remembered that these factors will no longer permit us to reproduce the correlations from their loadings (see page 22). Corresponding to these loadings, however, is another set which will reproduce the correlations. This is known as a *structure* and can be obtained by multiplying the original centroid loadings by the normalized inverse of the transformation matrix. We shall not calculate it because it is of no help in interpretation of the factors. The loadings which we have calculated are known as the factor *pattern* and it is from these that we are able to recognize the nature of the factors involved. This task we shall deal with in the next chapter.

We have now considered the methods of rotation which are likely to be most useful to a beginner. It remains merely to add a word of warning with regard to the use of the single plane method to adjust the factors from a cluster-directed analysis. Here the axes are so close to their true position that we can best begin by taking our original factor as our trial one, giving it a weight of 1·0. In this case we do not plot it against itself as this would lead to a diagonal set of points which would suggest a slope of 1·0 every time. This would completely distort the process.

Chapter VI

CLUSTER-DIRECTED ANALYSIS

Before proceeding further it will be useful to summarize the position as far as we have gone:

1. The factor method is based upon the correlation coefficient which is a mathematical statement of the degree of agreement between two series of measurements.

2. It is assumed that correlation is produced because similar influences are at work and it is these influences which we call factors. The extent to which the factor influences the result from a test is known as its loading for that test. If we know the factor loadings for two tests we know that they will correlate to the extent of the product of the two loadings.

3. The total correlation will be the sum of such products if there are several factors in common to the tests. We are then faced with the problem of deciding how much of the correlation is due to each factor and so the factor loading for each factor on the test.

4. One method of doing this is to calculate the maximum amount of correlation which can, mathematically, be attributed to one factor and then to repeat the process with the correlation coefficients so reduced. This is repeated until all significant correlation has been accounted for.

5. This, however, gives factors of constantly decreasing size and with an arbitrary reference frame. To make the factors likely to agree with functional entities we have to resort to rotation of the reference axes. This rotation is carried out in such a way as to increase the number of zero loadings and decrease the number of negative loadings, i.e. to attain 'simple structure'.

6. When physical measurements are handled in this way the resulting factors do seem to correspond to fundamental dimensions. We seem justified, therefore, in using the method as a means of discovering fundamental dimensions in new spheres.

7. In addition to the primary factors found in this way it is possible to calculate second order factors. When we analysed the intercorrelations of notepaper measurements in this way we found a general second order factor which corresponded to rectangular shape.

We have now to consider another approach to the factor problem. By sectioning the correlation matrix it is possible to calculate several factors at once and if the sections are based upon clusters of intercorrelations which all have relatively high values the group factors so obtained will approximate to simple structure. Such methods can effect a great saving in time.

In America the most widely used group method is that of Thurstone ('multiple group method') but it is normally used only as a basis for rotation. Thurstone makes only a rough selection of clusters and bases the position of the reference axes entirely on simple structure principles. With more care in selecting clusters, however, one gets factors which require very little rotation to bring them into simple structure position and it is possible to carry out the final rotation by purely arithmetical calculations which make the reference axes conform to the correlation clusters rather than to simple structure requirements. Usually such factors are very similar to what we should get by the simple structure methods except that they are orthogonal and instead of a second order factor we get an additional first order factor. For many purposes this is an easier solution to interpret. The additional factor will be a general one only if there really is something in common to all the tests. It corresponds to what Burt calls a 'basic' factor in connection with his group method which we describe in the next chapter.

The advantages of using a group-factor method based on correlation clusters may be summarized as below:

1. Several, indeed all, the factors may be calculated at once with a consequent saving in time.

2. If the factors are fully based on the clusters as in Burt's method, or the amended form of Thurstone's method which we suggest at the end of this chapter, no rotation is required.

3. It may be argued that, if the correlation clusters are well defined, the reference axes are more reliably determined than they would be by the simple structure methods.

4. The presence or absence of a general factor is obvious without second order analysis.

5. The factors are uncorrelated which makes for easier interpretation.

SELECTING THE CLUSTERS

Cluster-directed analysis depends upon the simple fact that tests which have a common factor can be expected to show consistent intercorrelation. This is the sort of situation which we found in our first correlation table (page 20). If only one factor is present the correlations must take this form or if one factor predominates there will be an approximation to this form.

Cattell[1] has argued that clusters of high correlations are likely to occur where two factors overlap. This is no doubt true but such a position can be recognized if the correlations within the cluster are arranged in order of size.

1. With a single factor dominating we get a simple arrangement with the largest correlations in the upper left corner and the smallest in the lower right corner:

x	·74	·65	·53	·47
·74	x	·54	·48	·40
·65	·54	x	·35	·29
·53	·48	·35	x	·18
·47	·40	·29	·18	x

2. With overlapping factors we get various modifications of this hierarchical pattern. If two factors are present they seldom run parallel with their loadings. The result is that, if the tests are arranged according to the correlations which are dominated by one factor, the correlations which are chiefly influenced by the other factor will break the hierarchy. The correlation cluster below shows a typical position:

	A	B	C	D	E	F
A	x	·46	·38	·28	·21	·15
B	·46	x	·35	·30	·28	·23
C	·38	·35	x	·28	·30	·27
D	·28	·30	·28	x	·27	·28
E	·21	·28	·30	·27	x	·36
F	·15	·23	·27	·28	·36	x

[1] R. B. Cattell, *Description and Measurement of Personality* (New York, World Book Co., 1946), p. 80.

Inspection of this matrix shows that the hierarchical order is broken after test C. It conveniently breaks down into two smaller clusters as shown below:

	A	B	C
A	x	·46	·38
B	·46	x	·35
C	·38	·35	x

	F	E	D
F	x	·36	·28
E	·36	x	·27
D	·28	·27	x

If the two factors have parallel loadings, however, there is no way in which they can be distinguished. They will simply come out as a single factor. Separation can be effected only by the addition of new tests to the battery. These new tests would have to be selected so that they would be likely to have high loadings on one of the factors to be separated and low loadings (preferably zero) on the other factor.

It should be emphasized that this difficulty of separating parallel factors is not one peculiar to cluster analysis. Factors can be separated only by the difference in their effects. In so far as they have the same effects they cannot be differentiated at all. But seldom do two factors have similar effects except within a very limited field. Speed and strength may substitute one for the other in tasks such as loading sand but speed is of little help in dragging a car from the mud nor will strength be of much help to the juggler.

Let us now consider how we can apply cluster-directed analysis to a simple example. We choose here a table of correlations obtained from a research into the nature of intelligence so that it will be evident how the method works out with empirical material. The following table contains the intercorrelations of ten tests administered to fifty third-year university students.

	1	2	3	4	5	6	7	8	9	10
1	x	47	20	38	44	20	60	00	07	26
2		x	00	−17	26	58	48	−28	20	20
3			x	69	41	20	05	52	20	−12
4				x	35	15	13	31	27	30
5					x	33	40	00	22	26
6						x	12	06	50	33
7							x	−15	08	28
8								x	06	−64
9									x	22
10										x

N.B.—Decimal points have been omitted for convenience and no figures have been inserted below the diagonal as these are simply reflections of those above the diagonal. The correlation of 1 with 2 is naturally the same as the correlation of 2 with 1. By working with only half the table we economize in several ways.

It will be seen that in this matrix there are several zero and near-zero correlations, which indicates at once that more than one factor is probably present. Our first task is to find a cluster of significant correlations. We locate the highest coefficient (·69). This we find to be the correlation between tests 3 and 4 so we now search for a third test which has high correlations with both these tests. Test 8 has the highest correlation with test 3 and it also has a high correlation with test 4 so we add this to our cluster.

We now look for another test which has significant correlation with the three already in the cluster. Test 5 seems the most likely since it has the highest correlation with test 3; but although it has a high correlation with test 4, it has zero correlation with test 8, so we have to abandon it. We have no better luck with any of the other tests. To check on this we write out all the correlations for the three cluster tests:

3	20	00	x	69	41	20	05	52	20	−12
4	38	−17	69	x	35	15	13	31	27	30
8	00	−28	52	31	00	06	−15	x	06	−64

It will be seen by simply inspecting the columns here that in every case we have either a near-zero or a negative correlation to spoil the hierarchy.

Inspection of this kind is often facilitated if the whole table is written out and the rows cut in strips. The appropriate strips can then easily be put together to check the desirability of admitting further tests to the cluster being considered. Working in this way one first selects the strips for the two tests having the highest correlation, places them together and then scans for two correlations of significant size in the same column. When such a pair of correlations has been found the strip for this new test is added to the other two and a search now made for a column of three correlations. If three of the right sign and appropriate size are found a further strip for this test is added and a further search begun.

The size of correlation which can be accepted in this way depends upon the probable error of the coefficients and their general level. If in doubt it is best to accept anything above ·10 and later to discard the test if it is found that it has higher correlations with another cluster. Usually one selects the cores of the most obvious clusters and when three or four clusters have been found in this way proceeds to allocate the remaining tests to whichever clusters they seem most related to.

Following this policy we assemble a second cluster with the correlation of tests 1 and 7 as its pivot. This cluster is more extensive than our first one. We assign tests 2 and 5 to it and test 10 provisionally. At this stage we have only tests 6 and 9 left. These correlate to the extent of ·50 and so can form the nucleus of a third cluster. Test 10 correlates more highly with this third cluster than with our second cluster so we transfer it.

Finding the Factors

To make clear the principle involved in this method of analysis let us consider our first cluster separately. We set out below the intercorrelations for this cluster together with the correlations of the cluster test with all other tests:

	3	4	8	6	9	10	1	7	2	5
3	x	69	52	20	20	−12	20	05	00	41
4	69	x	31	15	27	30	38	13	−17	35
8	52	31	x	6	6	−64	00	−15	−28	00

Before proceeding further we must insert the communalities in the diagonal cells of the cluster. In a small 3 × 3 matrix like this it will not be sufficient just to insert the highest correlation from the row or column. It will be necessary to calculate several times to make sure that the results are converging on the true values. The most effective way to begin will be by an estimate but until one has had sufficient experience to do this successfully it may be desirable to adopt a more pedestrian approach. Quite a useful procedure in a 3 × 3 table is simply to choose the highest correlation coefficient for the column with the highest total, the lowest for the column with the smallest total and the intermediate coefficient for the other column. Usually one finds that the communality for the column with the highest total should be rather higher than the highest coefficient in the column while for the column with the lowest total the communality is usually somewhat lower than the lowest correlation in the column. The intermediate columns should have communalities roughly proportional to their column totals when added without the communality. Knowing these facts one soon learns how to anticipate results and so is able to converge more rapidly on the true value, but in this example we shall simply take the three given coefficients and repeat them as communalities. These values are by no means correct but it will be seen how we gradually converge on the wanted values and how little it matters if we do make considerable error in selecting our trial values. We set out the working below:

	(69)	69	52	
	69	(52)	31	
	52	31	(31)	
Sums:	1·90	1·52	1·14	Grand total = 4·56 $2\cdot135^2$
Loadings:	·89	·71	·53	(dividing by 2·135)
Squares:	·79	·50	·28	(communality is square of loading)
Trial:	2·200	1·50	1·11	(column totals when new communalities are substituted for first estimate)
				Grand total = 4·61 = $2\cdot147^2$
Loadings:	·93	·70	·52	(dividing by root of grand total above).
Squares:	·86	·49	·27	

It will be seen at this stage that for two of our columns the new communalities from our second trial are only ·1 different from those of the first trial. We can therefore expect that another trial will give factor loadings which are reasonably close to the true values. This further trial will be involved in the next stage of our working with the original table so we now return to it.

We now take the communalities as found above and insert them in the diagonal cells of our cluster as originally set out. All the columns are now summed but we find the grand total for the first three columns only, those belonging to the cluster. We take the square root of this as before but we now divide it into *all the column totals*. In this way we find not only the first factor loadings for the tests in the cluster but the loadings of this factor on all the tests outside the cluster.

The reasoning behind this is really quite simple. If a test outside the cluster has a loading on the cluster factor it will show correlation with every one of the tests within the cluster and the sum of this correlation will be equal to the sum of the loadings of the tests within the cluster (which, remember, is itself the root of the grand total of the correlations within the cluster matrix) multiplied by the loading of the test concerned. A simple theoretical example will make this clear:

Loadings	·6	·5	·9	·7
·6	x	·30	·54	·42
·5		x	·45	·35
·9			x	·63
Sum 2·0				1·40

Sum of loadings of test inside cluster (theoretical) is 2·0
Sum of correlations in column outside cluster is 1·40
When we divide the latter by the former we get ·7 which is the theoretical loading from which we calculated this column.

It will be seen that we have a very powerful instrument in this approach. By selecting our clusters we are able to ensure that we are dealing with a factor which is a functional unit and not just an average of what is common to the whole battery of tests and we are able to calculate the loadings on this factor for all the tests concerned. In other words we have fixed our factor axes in advance by selecting our tests.

It has already been explained that the nature of the unrotated centroid factor depends on the average effect. Adding another test to the battery alters all the loadings for the other tests. If we have a battery consisting of three-quarters intelligence tests and one-quarter memory tests we shall get a general factor chiefly concerned with intelligence. If we reverse the proportions our general factor will be largely a memory factor (assuming for the moment that intelligence and memory are functional entities). It is to overcome this arbitrary fixing of the factor axes by the particular selection of tests which have been included in the battery that we resort to rotation. The need for rotation may be obviated by the simple process of putting together only tests which are dominated by a single factor. This is precisely what the cluster approach aims to do. Incidentally it reduces the amount of work by enabling the calculations to be carried out with only a part of the table instead of with the

It is to be noted, however, that when factors are calculated in this way they will be oblique or correlated factors unless the theoretical loading due to each factor is calculated and subtracted from the matrix of correlations before the next factor is calculated. Fortunately we can calculate several factors simultaneously without calculating residuals, and, by quite a simple procedure, discover to what extent our factors are correlated so that we may then rotate the factor axes to orthogonal positions. Such rotation is done by simple multiplication and does not involve any difficulties in deciding where the factor axes should go. This is decided by a mathematical calculation. This rotation is necessary before calculating the residuals from the factors obtained because of an important difference between oblique and orthogonal factors. With orthogonal factors we have simply to multiply the loadings together to get the correlations. This is not the case with oblique factor loadings. Here the correlation is not the product of the loadings because we have to allow for the fact that the reference axes are not at right angles. This becomes obvious when we remember that by oblique rotation we reduce many of our loadings to zero, without correspondingly pushing up loadings on another factor. With orthogonal rotation we cannot move one factor axis without correspondingly moving another.

We shall understand all this much better when we have dealt with our example. We set out below a detailed explanation of the steps involved. The working is given on pages 58–59 where our correlation table now appears sectioned according to the clusters we found.
whole.

1. The first step is to insert the communalities as we have explained in connection with our first cluster. We omit here the working out of these for the other two clusters. The reader can amuse himself by

attempting it. He will not get quite the same results as we have here because we used guessed values for our first trial in order to shorten the work. Further trials would, however, bring similar results and even a few trials should give results close to what we have here.

2. We now proceed to total each section separately and then to find the grand total for each section and set them out in a 3 × 3 table on the right. The diagonal entries here are those for our three clusters so we find their square roots and enter them in the corresponding table below.

3. We next proceed to divide our three rows of column totals by the corresponding root just found and carry on to repeat the same operation on the grand totals on the right. For the diagonal entry we shall, of course, get the figure already entered since dividing a number by its square root gives one the square root again.

We have now got the loadings for our first three factors and are ready to calculate the correlation between the factors.

4. The next step is to take each column of the 3 × 3 table on the right and divide each entry by the diagonal entry. The result is a new table in which all the diagonals are unity. The other entries in this table give us the information we are looking for. They tell us the correlation between our factors. We must spend a few minutes considering why this is so.

Consider the section I—II. The totals of the columns in this section are governed by the amount of correlation with factor I but obviously the absolute total depends upon how many coefficients are concerned. We therefore divide by the root of the cluster total in order to make allowance for this and we have to do the same to the grand total. But this grand total is also affected by the number of columns involved and some adjustment must be made for this too. The obvious adjustment is division by the root of the grand total of the cluster concerned. The tests concerned in section I—II are those of cluster II which give the correlations in section II—II. We therefore divide by the root of this cluster total. When this is done we have a figure which tells us the degree of correlation between factor I and factor II.

All this is involved in the double division by the cluster roots. When applied to the cluster totals themselves (the diagonal entries in the 3 × 3 matrix) it gives us unity indicating that the correlation of the cluster with itself is perfect.

5. We have now to find the multipliers which will enable us to get rid

Cluster-directed Analysis

	I			II			III				
	Ch	Ch Er	Obs	Class Er	Ass	Poets	70/1	Gott	70/23	P.D.	
Checking	(86)	69	52	20	20	−12	20	05	00	41	
Checking Errors		(49)	31	15	27	30	38	13	−17	35	I
Obsessional			(27)	06	06	−64	00	−15	−28	00	
Classifying Errors				(67)	50	33	20	12	58	33	
Assertive					(34)	22	07	08	20	22	II
Poets						(16)	26	28	20	26	
70/1							(67)	60	47	44	
Gottschaldt								(62)	48	40	III
70/23									(34)	26	
Progressive Designs										(28)	

Section totals	2·07	1·49	1·10	·41	·53	—·46	·58	·03	—·45	·76
	·28	·72	—·52	1·50	1·06	·71	·53	·48	·98	·81
	·66	·69	—·43	1·23	·57	1·00	2·18	2·10	1·55	1·38

Section grand totals

(4·66)	·48	·92
·48	(3·27)	2·80
·92	2·80	(7·21)

Rows divided by root of diagonals

Oblique factor loadings	·96	·69	·51	·19	·25	—·21	·27	·01	—·21	·35
	·15	·40	—·29	·83	·59	·39	·29	·27	·54	·45
	·25	·26	—·16	·46	·21	·37	·81	·78	·58	·51

2·16	·22	·43
·27	1·81	1·55
·34	1·04	2·69

Columns divided by root of diagonals

Correlations between factors	1·00	·12	·16
	·12	1·00	·58
	·16	·58	1·00

of this correlation between our factors. This involves two operations in matrix algebra which correspond to finding a square root and a reciprocal respectively in arithmetic. The reciprocal of a number is obtained by dividing that number into unity. The reciprocal of 2 is ½. Multiplying them together gives unity. The advantage of finding reciprocals is that multiplication by the reciprocal can be substituted for dividing by the original number and multiplication is often easier than division. The desired reciprocal can be found by consulting a table but the matrix equivalent, known as an inverse, must be calculated.

The first step, however, is to find the equivalent of the square root of the matrix. What we need is another matrix which, when multiplied by itself, gives us our original. A set of factor loadings does just this. They multiply out to produce the correlations as we have seen. What we need to do, therefore, is to find the factors in our matrix of factor intercorrelations. For this purpose Thurstone uses the diagonal method. We take as our first factor loadings the first column in the matrix. (This applies only when the diagonal is unity as here.) We then work out the correlations resulting from these loadings and subtract. The result is a 2×2 matrix because all the first row and column residuals are zero:

Factor I	1·00	·12	·16	Top row taken as factor loadings.
	1·00	·12	·16	Multiplying top row by 1.
1·00	00	00	00	Subtracting.
	·12	1·00	·58	
·12	·12	·01	·02	Multiplying top row by ·12.
	00	·99	·56	Subtracting.
·16	·16	·58	1·00	
	·16	·02	·03	Multiplying top row by ·16
	00	·56	·97	Subtracting.

This leaves:	·99	·56
	·56	·97

Our next factor can have only two loadings because our table has been reduced. This time we take the square root of the top diagonal entry and divide this into the second entry in the column to get the second loading. (The same thing was really involved with the preceding factor but as the square root of 1 is 1 and 1 divided into the other coefficients leaves them just the same there was no arithmetic to be performed)

Factor II

$$\begin{array}{lrrl} & \cdot 99 & \cdot 56 & \\ \sqrt{\cdot 99} = \cdot 99 \text{ (approx.)} & \cdot 99 & \cdot 56 & \text{Multiplying loadings by } \cdot 99. \\ \cline{2-3} & 00 & 00 & \text{Subtracting.} \\ \cline{2-3} \frac{\cdot 56}{\cdot 99} = \cdot 57 & \cdot 56 & \cdot 97 & \\ & \cdot 56 & \cdot 32 & \text{Multiplying loadings by } \cdot 57. \\ \cline{2-3} & 00 & \cdot 65 & \\ \cline{2-3} \end{array}$$

This leaves us with only ·65 for our last calculation and the root of this gives the loading. Putting our factor loadings into a table we have:

	1	2	3
1	1·00	00	00
2	·12	·99	00
3	·16	·57	·81

6. We must now find the inverse of this matrix. For tables beyond 3 × 3 this is somewhat complicated but it is quite simple for 2 × 2 or 3 × 3 tables with zeros above the diagonal. We shall here use the simplified method recently proposed by Fruchter[1] and the reader can consult Thurstone's text if he wishes to tackle anything more ambitious.

a. For the diagonal entries one simply inserts the reciprocals as found in any book of tables.

b. For the entry 2—1 (which is ·12 here) we take the fraction

$$\frac{2\text{—}1}{2\text{—}2} = \frac{\cdot 12}{\cdot 99} = \cdot 12 \text{ (approx.)}.$$

c. New entry for 3—1 becomes:

$$\frac{3\text{—}2 \times 2\text{—}1}{2\text{—}2 \times 3\text{—}3} - \frac{3\text{—}1}{3\text{—}3} = \frac{\cdot 57 \times \cdot 12}{\cdot 99 \times \cdot 81} - \frac{\cdot 19}{\cdot 81} = - \cdot 11.$$

---

[1] Fruchter, B., 'Note on the computation of the inverse of a triangle matrix', *Psychometrica*, 1949, 14, 89–94.

d. New entry for 3—2 becomes:

$$-\frac{3\text{—}2}{2\text{—}2 \times 3\text{—}3} = -\frac{\cdot 57}{\cdot 99 \times \cdot 81} = -\cdot 71.$$

So that our inverse is:

1	2	3
1·0	·00	·00
—·12	1·01	·00
—·11	—·71	1·24

7. We can now find our orthogonal loadings.

a. For factor I we multiply by the figures in our inverse above. That is, we multiply the original factor I loadings by 1 and the original factor II and III loadings by zero. In other words we make no alteration at all.

b. For factor II the second row in our inverse gives us the multipliers. Original II is multiplied by 1·01 (each loading) and to these figures are added original factor I multiplied by —·12. (The zero under 3, of course, makes no difference and is ignored.)

c. For factor III we have three sets of multiplication: original III by 1·23, I by —·11 and II by —·71.

We thus get:

	3	4	8	6	9	10	1	7	2	5
I′	·96	·69	·51	·19	·25	—·21	·27	·01	—·21	·35
II′	·04	·32	—·35	·82	·57	·42	·26	·27	·57	·41
III′	·09	—·04	—·05	—·04	—·19	·20	·76	·77	·36	·27

8. The next step is to calculate the amount of correlation that these three factors can account for and to subtract from the original table to see whether we have significant residuals left. As this is the same process that we described in connection with the simple centroid method we have not done it here. The only complication is that instead of subtracting one theoretical correlation from each of the originals we have the sum of three to subtract.

If the residuals still show a number of significant coefficients further analysis is carried out. In the case of this table there are only two residuals over ·3 and no consistent clusters so we assume there are no more factors of any importance to be considered.

So far we have been following Thurstone's method except that we have carefully selected our clusters for the sectioning of the matrix. There remains now to make a final adjustment. It will be seen that we have quite a fair approach to simple structure and with the large probable error of our correlation coefficients it might be difficult to improve upon this with any certainty, but the position of the reference axes is not quite right according to our cluster criterion.

The factors first calculated were oblique ones considerably distorted by overlap. We have rotated them to orthogonality (which, by the way, is necessary if we wish to calculate the theoretical amount of correlation accounted for by the factors) by a pivoting process. In effect we have left factor I in its original position and rotated factor II into a position at right angles to this. With these two factors fixed the third was rotated into a position at right angles to both. (This was involved in the diagonal analysis.) The effect of this process is to overweight the first factor and distort all subsequent factors. What we should have done was to make *all* the factors participate equally in the give and take required to attain orthogonality. In this way we should have caused the factor reference axes to pass as closely as possible through the centres of the clusters.

There are two ways in which we may bring about this result. Either we may adjust the results as calculated by the Thurstone method or we may substitute a different method for the diagonal analysis used by Thurstone and so calculate the cluster-based factors directly. The former we discuss in the appendix as it can be done only approximately. The latter can be carried out precisely by a method suggested by Dr Ledyard Tucker. It is based upon what is called the 'principal axes' method of analysis. This is a refinement of the centroid method which is rather cumbersome for use in most practical work but which we can use here because we have only a small inter-factor matrix and, by way of compensation, we can dispense with the calculation of an inverse.

The calculation of principal axes is very similar to the calculation of centroid factors. Burt refers to the latter as the method of 'simple summation' and the former as the method of 'weighted summation'. The essence of the method is that every row in the correlation matrix should be weighted by the loading for the corresponding variable. Of course, as soon as we do this the factor loadings change and the difficulty in the method is that we must continue the process until the values settle down and the amendment of the weighting makes no further difference to the factor loadings.

In practice we are able to shorten the work somewhat. We first add our columns in the usual way, omitting the grand total. We then divide our column totals by the largest one and use the figures so obtained as the weights for the corresponding rows. Having multiplied each row by its appropriate weight we again add columns and divide by the largest total. If the weights come out as before we are ready to calculate the factor loadings. Usually a number of repetitions will be required before this happens but after some experience the work can be considerably shortened by guessing the appropriate weights. Any guess will do but the nearer the guess is to the true weights the less the number of repetitions required.

When satisfactory weights have been obtained the column totals are proportional to the factor loadings, but we cannot use the usual centroid method to find the factor loadings since the weights we have used are not the factor loadings themselves but only proportional to them. We can get over this difficulty, however, because the sum of the squares of the factor loadings is equal to the sum of the loadings of the largest column. Knowing this we can find our factor loadings by first normalizing the column totals and then multiplying them by the square root of the highest column (the original value, not the normalized one). It may be useful to know that the highest column total is known as the latent root.

We set out the working for the first factor below. We have started with guessed weights rather close to the final weight in order to shorten the working:

Original matrix				Trial weights	
1·00	·12	·16		·44	·436
·12	1·00	·58		·98	·986
·16	·58	1·00		1·00	1·000
1·28	1·70	1·74	Column totals		
·440	·053	·070	multiplied by	·44	
·118	·980	·568	multiplied by	·98	
·160	·580	1·000	multiplied by	1·00	

·718	1·613	1·638	Column totals
·438	·985	1·000	Dividing by latent root (1·638)
·436	·052	·070	Multiplying by new weights. Note that these are not precisely those obtained from the column totals but involve a new guess at the final values.
·118	·986	·572	
·160	·580	1·000	
·714	1·618	1·642	Column totals
·435	·985	1·000	Dividing by latent root
·189	+·971	+1·000	= 2·160 Squaring and adding
·379	·859	·872	Multiplying by $\sqrt{\frac{1\cdot642}{2\cdot160}}$

It will be noticed that we have combined the operations of normalizing and multiplying by the root of the highest column. Where there are more variables the saving thus effected becomes important. It should be noticed also that this operation is performed on the *unsquared* numbers. The reader may be interested to take the factor loadings so obtained, use them as weights for the corresponding rows and then calculate an ordinary centroid solution. The result should be approximately similar loadings.

Second and third factors are found in a similar way from the residual matrices. We thus get:

I	II	III
·379	—·925	—·029
·859	·239	—·455
·872	·167	·462

We now multiply each latent root by its own square root and divide the columns of factor loadings by this figure:

Latent roots:	1·642	—·941	·420	
Square roots:	1·282	·970	·648	
Products:	2·105	·913	·272	
	·180	—1·013	·107	Columns divided by products above
	·408	·262	1·672	
	·413	·183	1·697	

To get our transformation matrix we now multiply this matrix by the original principal axis matrix above just as we did on page 47 when finding factor intercorrelation. In this case, however, we multiply row by row instead of column by column. The result is the matrix below:

I	II	III
1·01	—·04	—·06
—·04	1·17	—·37
—·06	—·37	1·18

It will be seen that whereas the Thurstone transformation matrix had no entries below the diagonal this matrix is symmetrical and so brings about corresponding adjustment in all the factor loadings when they are multiplied by it. We use it just as we did the corresponding Thurstone matrix and so we get the factor loadings below:

	Thurstone Rotation			Symmetrical Rotation		
Checking	**·96**	·04	·09	**·95**	·04	·18
Checking errors	**·69**	**·32**	—·04	**·67**	**·34**	·12
Obsessional	**·51**	**—·35**	—·05	**·54**	**—·30**	·11
Classification errors	·19	**·82**	—·04	·13	**·79**	·22
Assertive	·25	**·57**	—·19	·21	**·60**	·01
Poets (i min. list)	—·21	**·42**	·20	—·25	**·33**	**·31**
70/1 Intelligence	·27	·26	**·76**	·21	·03	**·83**
Gottschaldt figures	·01	·27	**·77**	—·06	·03	**·82**
71/23 Intelligence	—·21	**·57**	**·36**	—·27	**·42**	**·50**
Progressive designs	**·35**	**·41**	·27	**·30**	**·32**	**·41**

We have printed in bold face all factor loadings which reach or exceed ·30 to make easier a comparison with the Thurstone loadings. The highlights are not very different but there are more significant loadings for the third factor and two more near-zeros have appeared in the second factor loadings. If several more factors had been concerned probably there would have been much more substantial differences and it would have been dangerous to accept the Thurstone factors without rotation.

We now have reference axes based on our cluster analysis. Does the result appear satisfactory? We have several near-zeros but it should be remembered that we have a very high probable error in this material so that our zero equivalents may include loadings up to ·20 or even higher. Only one negative loading remains for factors II and III and it will be seen later that this is to be expected. There are three negatives for factor I but only two are of possible significance and a glance at the correlation matrix shows that the variables concerned have some negative correlations which justify such a relationship. In dealing with personality one can expect some negative loadings and they seem to make good sense here as will be seen when we consider their interpretation.

Chapter VII

COMBINING CENTROID AND GROUP ANALYSIS

Professor Burt has evolved an ingenious method of using group and centroid methods of factorization. This method, which seems little known outside Great Britain, has outstanding advantages for many types of factor problems. A centroid analysis of the complete matrix is first carried out in order to discover the main lines of division. On the basis of this the matrix is sectioned in much the same way as for cluster-directed analysis but before calculating the cluster-based factors Burt takes out what he calls a 'basic factor'. This is usually a general factor but not necessarily so since it is based, not on the whole matrix, but only on the inter-cluster correlation. We shall be in a better position to understand all this after considering a simple example. For this purpose we take the table of correlations from which we calculated the centroid factors used to illustrate rotational methods in chapter IV.

From this matrix, it will be remembered, we calculated three factors. To get three factors by the Burt Group Factor method it is sufficient to section into two parts since the basic factor will provide the third factor. We must note, however, that a matrix which yields only three factors by the centroid method may yield four factors if we take account of a basic factor. This happens when the group factors are correlated and the basic factor corresponds to this inter-factor correlation. We have already referred to the appearance of a general factor in the form of a second order factor when one uses the Thurstone technique. Burt's basic factor is a first order equivalent of this second order factor. Such a factor does not appear by the centroid method unless rotation is resorted to. With the present method we may check its presence by inspection of those sections of the correlation matrix which represent the inter-factor correlation. If the coefficients in any of these sections are consistently below the level of significance the corresponding cluster should be distributed among the other clusters. The reason for this will be seen from inspection of the detailed method.

Now let us proceed with our example. The second centroid factor divides our variables into four with positive and four with negative loadings. The latter are further subdivided by the third factor, variables 1 and 2 having negative loadings on both factors and variables 7 and 8 having negative loadings on factor II but positive loadings on factor III. If we thus section our correlation matrix into three parts we get:

1	2	3	4	5	6	7	8
x	·47	·42	30	01	20	12	18
	x	21	15	00	09	24	29
		x	42	29	33	−04	00
			x	35	33	−03	00
				x	28	−04	−02
					x	−03	−01
						x	31
							x

It will be seen that the correlation between the second and third clusters is insignificant. The highest correlation coefficient for this section is —·04. There is therefore little justification for seeking four factors in this case so we reduce our matrix to two sections which will give us three factors. These two sections we base on the division between positive and negative loadings on the first bipolar factor. We thus get the matrix produced below (p. 70) together with the column totals for the inter-cluster sections. It will be noticed again that we have not reproduced the correlations below the diagonal since these are just the mirror reflection of those above the diagonal. Rows in the upper right hand section thus correspond to columns in the lower left-hand section.

This matrix provides us with a very simple case which should make it easier to appreciate the principle involved. We may assume here that we have either two overlapping factors or two non-overlapping factors defined by the two clusters plus a third factor which spans both clusters. The latter is the assumption made by the present method and, since we know from our centroid analysis that at least three factors are present, it seems a justifiable assumption in the present case.

One thing we can be sure of: that the inter-cluster correlation must be due to something which variables 1, 2, 7 and 8 on the one hand, and variables 3, 4, 5 and 6 on the other hand, have in common. We proceed

1	2	7	8	3	4	5	6
x	47	12	18	42	30	01	20
	x	24	29	21	15	00	09
		x	31	−04	−03	−04	−03
			x	00	00	−02	−01
				x	42	29	33
					x	35	33
						x	28
							x
				59	42	−05	25
93	45	−14	−03				

on the assumption that we should account for as much as possible of this correlation by one factor common to both sets of variables.

We cannot carry out an ordinary centroid analysis of the section concerned because the coefficients in this case are not the result of multiplying one set of factor loadings by a similar set (pp. 57–60) but of multiplying one set by another *different* set (for the other cluster). We cannot, therefore take the square root of the grand total as our divisor for the column totals. Instead we must find two different divisors and our problem is to find the appropriate weighting for these. Burt solves the problem by resort to the original centroid analysis. He sums the first factor loadings for the two sets of variables. The ratio between these two totals gives the appropriate weights for the two divisors. For this example they are 1·72 and 1·88 respectively and from these we calculate as below:

$$\text{Divisor I} = \sqrt{\frac{1\cdot88}{1\cdot72} \times \frac{1\cdot21}{1}} = 1\cdot15$$

$$\text{Divisor II} = \sqrt{\frac{1\cdot72}{1\cdot88} \times \frac{1\cdot21}{1}} = 1\cdot05$$

Using these two divisors we find that the column totals yield the following factor loadings:

$$\underbrace{\begin{array}{cccc}1 & 2 & 7 & 8\\ \cdot81 & \cdot39 & -\cdot12 & -\cdot03\end{array}}_{\text{sum} = 1\cdot05} \qquad \underbrace{\begin{array}{cccc}3 & 4 & 5 & 6\\ \cdot56 & \cdot40 & -\cdot05 & \cdot24\end{array}}_{\text{sum} = 1\cdot15}$$

It will be seen that we check our calculation from the fact that the sum of one set of loadings is equal to the divisor used for the other set.

The rest of the work for this group analysis we need not show in detail as it involves nothing new. We first calculate the theoretical correlation which would be accounted for by this basic factor (remember that, although we calculated from two sections, this is one factor and not two) subtract this from the original correlation coefficients and then proceed to make a centroid analysis of the two residual clusters, i.e., the diagonal sections of the matrix. There are now no significant correlation coefficients outside these sections so we have two non-overlapping group factors:

	1	2	7	8
Factor II	·35	·50	·58	·58

	3	4	5	6
Factor III	·44	·51	·70	·43

For convenience of comparison we set out the loadings on the three factors with the variables in their original order.

	I	II	III
1	·81	·35	
2	·39	·50	
3	·56		·44
4	·40		·51
5	—·05		·70
6	·24		·43
7	—·12	·58	
8	—·03	·58	

It will be seen that despite the difference in method, which has reversed the position of the second and third factors, the general pattern of loadings is remarkably similar to our rotated centroid solution.

The chief difference is that the factors are now definitely orthogonal and there is no overlap except for the basic factor. This absence of overlap can usually be regarded as an artificial result. We can usually expect some degree of overlap with empirical material and to take care of this situation Burt resorts to an ingenious artifice. He rotates the original centroid factors just as Thurstone does but instead of proceeding according to the simple structure scheme he bases his rotation on the group analysis. Firstly, in order to simplify the work he condenses the table of factor loadings by summing the loadings for each section. For our example this becomes:

	I	II	III
A	1·20	·85	
B	—·15	1·16	
C	1·15		2·08

These condensed loadings are now used to calculate a table of theoretical correlations as before, the correlations resulting from each factor being added together. We thus get the following table of inter-section correlations:

A	2·16	·81	1·38
B		1·37	—·17
C			5·65

Since this is a theoretical table we have the diagonal entries included and can make an exact centroid analysis of it. Working to two decimal places we get:

I	II	III
1·20	·61	·60
·55	·84	—·60
1·89	—1·44	

We now have three group factors and three centroid factors which are, within rounding errors, equivalent. By calculating a transformation matrix which will convert one into the other we acquire the means to transform our original centroid factor loadings (before condensation) into an equivalent basic and group factor solution which will now admit any overlap which may be required. We give below the working required in the calculation of the transformation matrix:

Group factors			Centroid factors		
I	II	III	I	II	III
1·20	·85		1·20	·61	−·60
−·15	1·16		·55	·84	·60
1·15		2·08	1·89	−1·44	

Let us consider factor I first. What we need is a set of multipliers A, B, C, such that when the items in Group Factor I column are multiplied by A, the items in Group Factor II column by B and the items in Group Factor III column by C the sums of these products will be the items in Centroid Factor I column. That is:

$$\begin{aligned}(1\cdot20 \times A) + (\cdot61 \times B) + (-\cdot60 \times C) &= 1\cdot20\\ (-\cdot55 \times A) + (\cdot84 \times B) + (\cdot60 \times C) &= 1\cdot55\\ (1\cdot89 \times A) + (-1\cdot44 \times B) &= 1\cdot89\end{aligned}$$

Now fortunately there is a principle of matrix algebra which enables us to simplify our problem somewhat. If we have three matrices, C, T and G such that

$$C \times T = G$$

then $G \times T' = C$ (T′ being the transpose of T, i.e. with columns and rows interchanged)

Taking advantage of this principle we can set out a simpler series of equations based upon the reverse process of transforming the Group Factor matrix into the Centroid Factor matrix and, after calculating this transformation matrix from the equations, just interchange columns and rows and so get the transformation matrix we really require.

Factor I. We have:

$$\begin{aligned}1\cdot20\,A + \cdot85\,B &= 1\cdot20 &&\text{(i)}\\ -\cdot15\,A + 1\cdot16\,B &= \cdot55 &&\text{(ii)}\\ 1\cdot15\,A + 2\cdot08\,C &= 1\cdot89 &&\text{(iii)}\end{aligned}$$

so $1\cdot64\,A + 1\cdot16\,B = 1\cdot64 \left(\text{line (i) multiplied by } \frac{1\cdot16}{\cdot85}\right)$

$$\text{and } 1\cdot79 = 1\cdot09 \text{ (subtracting line (ii))}$$
$$A = \cdot61$$

We can now find the true value of A in equation (i) and so calculate the value of B:

$$\begin{aligned} \cdot 73 + \cdot 85\,B &= 1\cdot 20 \\ \text{so}\quad \cdot 85\,B &= 1\cdot 20 - \cdot 73 \\ \text{and}\quad B &= \cdot 55 \end{aligned}$$

In the same way we can calculate the value of C by substituting in equation (iii):

$$\begin{aligned} \cdot 70 + 2\cdot 08\,C &= 1\cdot 89 \\ \text{so}\quad 2\cdot 08\,C &= 1\cdot 89 - \cdot 70 \\ \text{and}\quad C &= \cdot 57 \end{aligned}$$

Factors II and III. An exactly similar procedure is required to make the calculation for these two factors. We merely use the entries from the Centroid Factor II or III column in place of those from the Centroid Factor I column.

We thus finally get the following three columns for our transformation matrix:

I	II	III
·61	—·01	—·79
·55	·73	·41
·57	—·69	·44

Transposing columns and rows this becomes:

I	II	III
·61	·55	·57
—·01	·73	·69
—·79	·41	·44

We can check the rows of this (original columns) by finding the sums of the squares of the entries. These sums, apart from rounding errors, should be unity. In this case we find they are:

I	II	III
·3721	·0001	·6241
·3025	·5329	·1681
·3249	·4761	·1936
·9995	1·0091	·9858

These are sufficiently close approximations to unity for two-figure working.

We now use this transformation matrix to rotate our original centroid factors and we thus obtain the following orthogonal rotation which will be seen to be quite similar to our slightly oblique simple-structure rotation:

	I	II	III
1	·79	·01	·35
2	·37	—·01	·51
3	·52	·45	·04
4	·36	·54	·05
5	·01	·65	·00
6	·24	·45	·01
7	—·07	—·05	·53
8	—·02	—·03	·57

This example shows how a three factor table, i.e. a basic factor and two group factors, is handled. We must now explain how the divisors for the column totals are found when there are more than three group factors.

Three Group Factors. The matrix is sectioned into nine parts which we shall represent by the following letter scheme:

$$\begin{array}{ccc} \text{X} & \text{A} & \text{B} \\ \text{A}' & \text{Y} & \text{C} \\ \text{B}' & \text{C}' & \text{Z} \end{array}$$

Each section is summed separately and then grand totals for each column are obtained but with the diagonal sections X, Y, Z omitted. It is these column totals which are to be divided.

For the A′, B′ columns the divisor is:

$$\sqrt{\text{C}}\left\{\frac{\text{A}+\text{B}}{\sqrt{\text{AB}}}\right\}$$

For A, C′ columns the divisor is:

$$\sqrt{\text{B}}\left\{\frac{\text{A}+\text{C}}{\sqrt{\text{AC}}}\right\}$$

For B, C columns the divisor is:

$$\sqrt{\text{A}}\left\{\frac{\text{B}+\text{C}}{\sqrt{\text{BC}}}\right\}$$

It should be noted that, so far as section totals are concerned the three sections A, B, C correspond exactly with A′, B′, C′ so that the latter letters are not used in the formula. Notice also that the letters within the braces always give a clue as to the column totals concerned.

More Than Three Group Factors. Under these conditions the calculation of the column divisors is based upon a centroid analysis of the section totals (with the diagonal sections again omitted). This analysis must be repeated several times until the communalities are equal to the squares of the loadings obtained. Each of these loadings is now subtracted from the sum of the loadings multiplied by the number of group factors less one. This process yields the set of divisors required and if no errors have been made it will be found that the sum of these divisors is identical with the figure from which the factor loadings were subtracted to give these divisors.

MERITS OF THE METHOD

1. As a group-factor method it provides a useful method for the selection of clusters. It will have been noticed that by sectioning according to the divisions suggested by the centroid analysis we have obtained, in the example used here, a solution which corresponds to that obtained by seeking simple structure. It should be noted also that the division into nine sections which we first tried out and rejected because of the absence of significant correlation in one section, is just the arrangement at which one would arrive for a cluster-directed analysis and would give a cluster-directed solution very similar to what we obtained by Burt's method. The reason for this is that Burt's method gives one extra factor: the basic factor. In the cluster-directed method this has to be accounted for by another cluster.

2. It provides a solution in terms of orthogonal factors with the basic factor replacing a second order factor. It will be remembered that in our example the oblique factors showed negligible intercorrelations.

Chapter VIII

INTERPRETING THE RESULTS

We have now considered how we can resolve a set of correlations into several factors and ensure that the factors have reference axes which are likely to correspond with causal entities. If we have done this well the most important step in interpretation has already been performed but the rest is by no means automatic.

The basis of all interpretation is the assumption that tests which have high loadings on a common factor have something in common. Just what this is cannot always be judged with certainty. If our research has not been well planned we may be quite unable to decide. But it may be possible to advance several hypotheses which can be tried out in a further investigation. Factorial research depends just as much upon the formulation and testing of hypotheses as any other form of research.

As a practical example we can now consider the results we obtained in our cluster-directed analysis. The material for this analysis is by no means easy to handle and we chose to use it so that we should not give a false conception of the power of factor analysis. The correlations have a probable error of about ·15, which means that the pattern may be considerably distorted by chance factors. Furthermore, the reliability of many of the tests is low because of their shortness and this introduces a further chance variation which tends to depress some of the correlation coefficients.

Let us consider to what extent we have a meaningful pattern emerging despite these distorting influences. To facilitate this we collect together all the high loadings for our three factors and list them in order of size.

Factor I		Factor II		Factor III	
Checking	·95	Classif. Errors	·79	70/1	·83
Checking Errors	·67	Assertive	·60	Gottschaldt Fig.	·82
Obsessional	·54	70/23	·42	70/23	·50
Progressive Design	·30	Check Errors	·34	Progress. Des.	·41
		Poets	·33	Poets	·31
		Progress. Des.	·32		
		Obsessional	—·30		

We now proceed upon the principle that tests which get loadings on the same factor are sharing in a common effect and that the size of the factor loading is an indication of the degree to which they are exerting this effect. We must bear in mind, however, that low reliability depresses the general level of correlation and so reduces the factor loading. Ratings of personality qualities usually have very low reliability and this is especially true of self-ratings. The loading for obsessional tendency in factor I may well, therefore, be unduly low. In this case we may suspect that the factor is largely one of temperament rather than one of ability. We systematically note this and other points of relevance:

1. Obsessional temperament, allowing for reliability, is highly important.
2. The factor is manifested mainly in speed of checking.
3. Whatever increases this speed also increases the number of errors.
4. More complicated tests are not significantly affected.
5. It has been shown that people of obsessional type tend to get very high scores in some simple tasks such as dotting moving circles.[1]

In the light of these observations we are inclined to regard this factor as one of obsessional drive rather than one of ability but it is quite possible that there may be some overlap with an ability factor. It would be necessary to design a further investigation to check this.

It will be remembered that we found two small but possibly significant negative loadings for this factor. The variables concerned were 70/23, *Intelligence* and *Poets*. If we are correct in interpreting this factor as obsessional drive these two variables should have loadings similar in sign to their correlation with the variable *Obsessional tendency*. We find these correlations to be —·28 and —·64 respectively and so we conclude that these small negative loadings for these variables further confirm our interpretation.

We may now consider our second factor. Here we make the following observations:

1. The highest loading is concerned with the number of errors made.
2. A temperament quality is again of major importance. Possibly its loading of ·60 is just as important as the higher loading for the more objective type of test.
3. Checking Errors also has a small loading on this factor.
4. Obsessional Tendency appears with a small negative loading which suggests that the factors are in some way contrasted although they both produce errors.

---

[1] Culpin, M., and Smith, M., *The Nervous Temperament*, Industrial Health Research Board Report No. 61 (London, H.M.S.O., 1930).

5. The Progressive Designs test is one which many subjects find very discouraging since they may wrestle with some of the items for several minutes without getting any clue as to the solution.

We are led to conclude that again we have a factor primarily concerned with drive but one which produces error chiefly in the more complicated material, probably because of over-confidence. Assertive people may be regarded as impulsive and lacking in caution. By contrast the obsessional type is dogged but usually very attentive to detail and this would account for the small negative loading here. So far as our data guide us we may, therefore, regard this factor as one of confident ambition. (It is of interest to note that, in another investigation, we have found this factor to be related to examination success.)

Our third factor presents much more difficulty and it is here that we realize how subjective factor analysis can be. The statistical results in themselves simply constitute a challenge and the best we can do is to formulate some hypothesis which can be verified or disproved by further work. What have we to guide us here?

1. Our two intelligence tests get high loadings but we note that the simpler test (70/1) gets much the higher loading.

2. The Gottschaldt test has a loading almost as high as that for 70/1 which suggests that it is not simply a factor of general intelligence which is involved.

3. Our first four tests all involve visual designs and so may involve a 'spatial' factor.

4. The highest loading being on the Gottschaldt test suggests that the factor may be one of plasticity since this test involves the ability to discover 'hidden' figures. One has to overcome the obvious appearance to see the lines forming a less obvious design. It is possible that in all these tests the ability to flit lightly from one possibility to another is likely to be helpful. The mind that easily gets into a rut will be handicapped.

5. Another possibility to be considered is that there is involved here something in the nature of span of apprehension. The mind which can attack on the broadest front will have an advantage over the one which has to work piecemeal over the material.

It will be obvious that, from these factor loadings, we can come to no definite conclusion as to what the factor is. We can only consider likely hypotheses and design further experiments to put them to the test.

This particular example illustrates well the extent to which factorial technique is dependent upon other than mathematical principles. Only a person intimate with the science concerned can plan the investigation or interpret the results. There is still the same scope for scientific imagination, the same possibility for genius to substitute for laborious plodding.

Chapter IX

THE ROLE OF FACTOR ANALYSIS

We have considered the fundamental principles of factor analysis and some of the chief ways in which the technique can be applied. It remains now to evaluate the role which factor analysis can play as a scientific method. This we shall do mainly with reference to psychology since this is the only field with which the author can really claim acquaintance. Students of other sciences will be able to make the appropriate translation to their own fields of study.

Psychology has a special claim on factor analysis since it is in this field that most of its development has taken place. If necessity is the mother of invention we can easily understand the reason for this, for it would be difficult to find any other science which has greater need for a technique of this kind. The fundamental method of science is to look for concomitance of variance. We observe what happens with what. We do something and note the results. Provided that we can vary one thing at a time this technique is a powerful one. We drop two weights of different amounts from the same tower at the same time and note the speed at which they fall. If they both reach the ground together we conclude that difference in weight does not influence the speed of falling bodies, but if we are to draw this conclusion it is essential that they differ in respect to weight only. In all other respects, e.g. wind resistance, both bodies must be given identical conditions. It is not always easy to secure such conditions, but in the physical sciences it is very often possible. In a science such as psychology it is very seldom possible. Most of the things which really matter about human behaviour involve variables which cannot be isolated. In the field of psychological measurement intelligence holds a unique place but there is no means by which we can vary intelligence in an experiment without changing other variables. We may choose subjects who undoubtedly vary in intelligence, but they probably differ in a thousand other ways too. We may use a test which is thought to be a measure of intelligence, only to find that it is influenced by other factors also as we have just seen. Hunger, headache, fatigue, bad temper, a broken heart, or ingrowing toe nails, may affect the results. For the psychologist it is essential to study many variables at the same time.

This multiplicity of variables makes it extremely difficult to note what goes with what, and the position is further complicated by the fact that the amount of variance due to any one factor may be quite small so that in the single case it becomes lost in mere chance effects. It is here that factor analysis comes to our aid. It enables us to base our observations on a large number of cases, with the vast increase in certainty which this makes possible, and at the same time to disentangle the complicated causal net. In this respect it is a most powerful tool and one which will be resorted to with greater frequency as the tedium of arithmetical labour becomes reduced by the development of computing machinery. The time will undoubtedly come when the arithmetical part of the procedure will be relegated to computing devices and the factor analyst will simply have to plan the investigation and interpret the results. But that time is not yet.

LESS BIAS OF POPULAR OPINION

Much experimentation in psychology has been based upon popular opinion. Memory, imagination, fear and even intelligence are popular concepts. If the psychologist takes them over and simply attempts to define them more rigorously he may be led into false paths. Instead of relying on popular opinion in this way for his concepts the factor analyst may look for functional units on the basis of his analysis and then consider how to name them appropriately.

This is rather important in the field of testing. There has been a considerable tendency to design tests empirically for a given purpose without attention to their basic nature. A test found to be effective in the selection of clerical workers is regarded as a clerical test and it proves somewhat disconcerting when the same test turns out to be useful also in selecting engineers. A factor analysis brings out the basic factors and gives us greater understanding both of why the test works in its original field and what other fields may find it useful. A test developed to measure speed of decision in aviation applicants was found to involve three distinct factors: perceptual speed, the ability to handle spatial relations, and psychomotor precision.

Exploration by the factorial method may be expected to open up new veins of thought just because of this fact. Although it is dependent on the formulation of hypotheses just as is any other scientific technique it does provide a more helpful framework.

ECONOMY OF TESTS

If we can base our psychological tests on factor studies the number of tests required should be immensely reduced. The number of major factors concerned in the chief human activities is likely to be much less than the number of activities. If we can develop factorially pure tests we shall have a psychological alphabet which will enable us to spell out an appropriate battery for any given purpose. Some psychologists will no doubt decry such an 'atomistic' approach. We can only reply to them that such an objection, carried to its logical conclusion, would abandon tests anyway. Any sort of testing assumes functional units of some kind.

VALIDATION OF TESTS

A unique function of factor analysis is concerned with the validation of tests. The best known tests in psychology are those concerned with intelligence. In order to justify intelligence tests it was necessary to show the degree of efficiency with which they measured intelligence. For this purpose there had to be some criterion. A common practice was to compare the results of the tests with the ratings given by teachers. But teachers' ratings are usually criticized because of their unreliability, and the chief reason for developing intelligence tests was to find a more efficient means of measuring intelligence. Are we then to justify the supposedly better because it agrees well with the supposedly worse? It seems absurd, and yet there is no obvious alternative. If we can find a better criterion of intelligence than the test we shall not need the test!

It is here that factor analysis comes to our rescue. We can correlate our tests with teachers' ratings and analyze all our correlations. If the tests and ratings have loadings on a common factor we can conclude that they are measuring the same thing *but the degree to which they measure will be shown by the size of the factor loading*, and this may well be higher for the tests than for the ratings. In this way we use our ratings simply to identify the factor, but we rely on other means to assess actual efficiency.

THE STATUS OF FACTORS

Just what are these things called factors? It has been asserted on the one hand that they are simply useful systems of classification and on the other that they correspond to real psychological or other units. We may agree with both statements if we are prepared to admit that there are degrees of reality. Whether or not we can usefully conceive of degrees of reality we can safely admit that there are degrees of certainty about reality. We analyze our world into units according to function, and the degree to which we find a thing real may be considered to vary with the evidence we have of its functional unity. I regard my hand as real because in endless experience I find it functioning as a consistent unit. Factor analysis is simply a means by which I observe functional unity in a more difficult field, and effectively applied it should reveal units with similar consistency. Such units must turn out to be the same units that we should find by other scientific means.

An analysis of visual ability should reveal some factors which correspond to certain variables in connection with the structure of the eyes, as for instance, range of accommodation or astigmatism. The observed functional unity is as much evidence of reality in the one case as the other.

The doubt which has been cast upon the 'reality' of factors is largely due to the use of forms of analysis which pay no attention to the position of factor axes. If the factor axes are determined simply by an arbitrary selection of tests and the first factor gets the maximum weight from the averaging process we can expect that factors will vary widely from one investigation to another and we have *no evidence of functional unity*. Such factors cannot be regarded as real. But if we find in repeated investigations the same functional units we may conclude that we are measuring something stable which may be rightly regarded as just as real as the study table which I repeatedly find in the same form.

APPENDIX

BASING FACTORS ON CORRELATION CLUSTERS

The method described in chapter V for locating factor reference axes according to the correlation clusters may be shortened by using an approximation method. In this case the analysis is first completed according to the directions given for the original Thurstone multiple group method and then the factors multiplied by a correction matrix. This correction matrix is built up cell by cell so that when the original transformation matrix is multiplied by it the product is a symmetrical matrix similar to the one obtained by the method suggested by Ledyard Tucker and used in the text.

This process has to be carried out in two stages. First we build a matrix with unity in the diagonal cells. This enables us to make our calculations satisfactorily. We then normalize the matrix thus obtained. This last step slightly distorts our matrix but still leaves us with a fair approximation to a symmetrical product.

With a 3 × 3 matrix as in the text there are only three items to be calculated for the first stage of the work since the diagonals are unity and the entries below the diagonal are the negatives of the corresponding entries above. If we refer to the cells of our matrix by subscripts according to the row and column concerned ($_{23}$ means second row, third column) the required items are obtained as follows:

$$c_{23} = \frac{-t_{23}}{t_{22} + t_{33}}$$

(where c refers to the cell in the correction matrix and t to the cell in the original transformation matrix)

$$c_{13} = \frac{-(t_{13} + t_{12} \times c_{23})}{t_{11} + t_{33}}$$

$$c_{12} = \frac{t_{13} \times c_{23} - t_{12} - t_{23} \times c_{13}}{t_{11} + t_{22}}$$

To anyone unfamiliar with the notation these formulas may seem rather imposing but they merely indicate what to multiply and add together and are really little more complicated than a knitting pattern. We apply them to the original transformation matrix below:

$$\begin{array}{ccc} 1{\cdot}00 & -{\cdot}12 & -{\cdot}11 \\ & 1{\cdot}01 & -{\cdot}71 \\ & & 1{\cdot}23 \end{array}$$

$$c_{23} = \frac{-(-\cdot71)}{1\cdot01 + 1\cdot23} = \cdot317$$

$$c_{13} = \frac{-(-\cdot11 - \cdot12 \times \cdot317)}{1\cdot00 + 1\cdot23} = \cdot066$$

$$c_{12} = \frac{-\cdot11 \times \cdot317 + \cdot12 + \cdot71 \times \cdot066}{1\cdot00 + 1\cdot01} = \cdot066$$

So our matrix is

$$\begin{array}{ccc} 1\cdot000 & \cdot066 & \cdot066 \\ -\cdot066 & 1\cdot000 & \cdot317 \\ -\cdot066 & -\cdot317 & 1\cdot000 \end{array}$$

and when normalized this becomes

$$\begin{array}{ccc} \cdot991 & \cdot063 & \cdot063 \\ -\cdot065 & -\cdot951 & \cdot302 \\ -\cdot065 & -\cdot302 & \cdot951 \end{array}$$

This is a much quicker process than calculating principal components and probably gives a sufficiently close approximation to ensure correct interpretation of factors. In this case the results are very similar. Below we give the precise symmetrical matrix which we obtained by the Tucker method together with the matrix we obtain by multiplying the Thurstone matrix by the correcting matrix obtained by this method:

$$\begin{array}{ccc} & \text{Precise} & \\ 1\cdot01 & -\cdot04 & -\cdot06 \\ -\cdot04 & 1\cdot17 & -\cdot37 \\ -\cdot06 & -\cdot37 & 1\cdot18 \end{array} \qquad \begin{array}{ccc} & \text{Approximate} & \\ 1\cdot01 & -\cdot02 & -\cdot08 \\ -\cdot02 & 1\cdot17 & -\cdot37 \\ -\cdot08 & -\cdot37 & 1\cdot18 \end{array}$$

It will be seen that these two rotation matrices would give factor loadings which would hardly differ within the second decimal figure.

For two-figure working the correction matrix may be rounded to two figures but it may be advisable to first calculate it to three figures as we have done above so as to approach as closely as possible to the correct figures. Remember that this correction matrix is applied to the orthogonal factor loadings produced by the Thurstone method and not to the original oblique factors as was the case when we employed Tucker's solution. It is possible to operate directly on the oblique factors, however, by first multiplying the Thurstone transformation matrix by our correction matrix as we did above for comparison with Tucker's precise solution. In this case it should be remembered that, since the results are only approximate, the correlation residuals will also be only approximate. This may be important if an attempt is made to calculate further factors from them.

For those who wish to try this approximation method we supply formulas to cover a 5 × 5 matrix. They can be adjusted for use with any smaller matrix by a proportionate reduction of all the subscripts. If the matrix is 4 × 4 all the 5's in the subscripts become 4's, all 3's become 2's and so on, the 1's just disappearing. They appear formidable but, as will be seen from the example just worked, the calculations involved are simple, and it is chiefly a matter of watching the decimal point and seeing that the right sign is inserted.

FORMULAS FOR CORRECTION MATRIX

The formulas given below apply to a 5 × 5 matrix. For use with smaller tables all subscripts must be proportionately reduced. For a 3 × 3 matrix the first formula becomes reduced to c^{23} etc. and the last seven formulas are ignored.

$$c_{45} = \frac{-t_{45}}{t_{44} + t_{55}}$$

$$c_{35} = \frac{-(t_{35} + t_{34}c_{45})}{t_{33} + t_{55}}$$

$$c_{34} = \frac{t_{35}c_{45} - t_{34} - t_{45}c_{35}}{t_{33} + t_{44}}$$

$$c_{25} = \frac{-(t_{25} + t_{24}c_{45} + t_{23}c_{35})}{t_{22} + t_{55}}$$

$$c_{24} = \frac{t_{25}c_{45} - t_{24} - t_{23}c_{34} - t_{45}c_{25}}{t_{22} + t_{44}}$$

$$c_{15} = \frac{-(t_{25} + t_{14}c_{45} + t_{13}c_{35} + t_{12}c_{25})}{t_{11} + t_{55}}$$

$$c_{14} = \frac{t_{15}c_{45} - t_{14} - t_{12}c_{24} - t_{13}c_{34} - t_{45}c_{15}}{t_{11} + t_{44}}$$

$$c_{23} = \frac{-t_{35}c_{25} - t_{34}c_{24} + t_{25}c_{35} + t_{24}c_{34} - t_{23}}{t_{22} + t_{33}}$$

$$c_{13} = \frac{-t_{35}c_{15} - t_{34}c_{14} + t_{15}c_{35} + t_{14}c_{34} - t_{13} - t_{12}c_{23}}{t_{11} + t_{33}}$$

$$c_{12} = \frac{-t_{25}c_{15} - t_{24}c_{14} - t_{23}c_{13} + t_{15}c_{25} + t_{14}c_{24} + t_{13}c_{23} - t_{12}}{t_{11} + t_{22}}$$

FURTHER READING

1. G. H. Thomson, *The Factorial Analysis of Human Ability*, 4th ed. (London, University of London Press, 1950).

A suitable text for readers with little mathematical background.

2. C. Burt, *The Factors of the Mind* (London, University of London Press, 1940).

Contains material not available in other texts and should be consulted by all advanced students.

3. L. L. Thurstone, *Multiple-Factor Analysis* (Chicago, University of Chicago Press, 1947).

The most comprehensive treatment of the subject yet offered but unfortunately it neglects the British developments. Essential as a book of reference and for all who wish to master the underlying mathematical principles.

Melbourne University Press, Carlton, N.3, Victoria

6656

POST-GRADUATE BASIC TRAINING COLLEGE
सत्यमेव जयते
Banipur, Po. Baigachi, 24-Pgs.

MELBOURNE UNIVERSITY PRESS
CARLTON, N.3, VICTORIA, AUSTRALIA
London and New York: Cambridge University Press